Return to Education and Education Mismatch in Indonesia

Widdi Mugijayani

Economic Section

Cardiff Business School

Cardiff University

A Thesis Submitted in Fulfilment of the Requirements for the Degree of Doctor of Philosophy of Cardiff University

March 2020

I would like to dedicate this thesis to my beloved parents, husband, children, and family.

Acknowledgements

I would like to express my sincere gratitude to my supervisors Professor Melanie Jones, Professor James Foreman-Peck and Dr. Ezgi Kaya, who gave me a lot of supports, motivations, feedbacks as well as guidance to finish this thesis. I have learned a lot for the last three years. And also, my sincere gratitude to Professor Partick Minford for the supports and help since PhD registration process. I would also like to thank to Dr Samuli Leppälä, the review panel convenor of this thesis. And I am most grateful to Lembaga Pengelolaan Dana Pendidikan (LPDP), the Republic of Indonesia, for providing me scholarship, this thesis would not have been possible without their supports.

Summary

This thesis aims to examine the effects of education expansion on labour market outcomes in the waged sector in Indonesia between 2000 and 2014 (the conditions before and after the education reform) using the Indonesia Family Life Survey (IFLS) data.

Chapter 3 finds the return to education increases with education level based on Mincer wage equation with OLS model. However, the return to education tends to decline for most education levels during the period of education expansion, consistent with an increase in the supply of educated labour. By sector, the return to education is generally higher in the public sector than in the private sector, possibly due to the lack of competition in the former. The findings are robust to accounting for the endogeneity of education and selection into the labour market.

Chapter 4 analyses both undereducation and overeducation resulting from education mismatch. It is found that the mismatch increases with the main driver being the increase of undereducation between the periods, based on Realised Match (RM) – mode and mean. This study also finds that education mismatch is determined by age, sector and firm size based on the Multinomial Logit (MNL) model; consistent with the Assignment Models. Yet, the determinants are sensitive to the different methods used, the sector/gender and the periods.

Chapter 5 finds both overeducation and undereducation incidences affect the labour wages, based on the Overeducation–Required–Undereducation (ORU) model (Duncan and Hoffman, 1981) with several panel methods. The return to one year of surplus schooling is the same with the return to an additional year of required schooling, and the return to one year of deficit schooling is negative and significant with the absolute value lower than the return to an additional year of required schooling. The results are sensitive to accounting for unobserved individual heterogeneity (fixed effects).

Table of Contents

List of Figures

List of Tables

List of Notations

A	Area/regional	PV	Present value
B	A person specific constant	R	Market interest rate
C	Cost of education	REQ	Required years of schooling
D	Demand	R_x	The internal rate returns
e	Exponential value	r	Market rate of interest
E	Equilibrium	S	Supply
Es	Separating Equilibrium	t	Time; t=1, 2, 3, … T-1
F	Work related and firm size	T	Maximum time
g	Growth	u	Solution in human capital derivation
HH	Household characteristics	UE	Undereducated / years of deficit schooling
h	Human capital investment	Uni	University
H	Human capital fraction	V	Part of IMR calculation (function of alpha)
Inv	Investment in schooling	v	Flow rate of mortality
i	Individual (i=1…I)	W	Wage
K	Explanatory variables	Ω_I	Proportion of group I
L	Labour market experience	WF	Workforce
l	Employment/labour/relative employment	X	Schooling, education level
M	Match	$\bar{X}$	Mean value of schooling
m	Number of matches	$\hat{X}$	Predicted value of schooling
MC	Marginal cost of schooling	Y	Non-Schooling
n	Number of variables	Z	Instrument variable
OE	Overeducated / years of surplus schooling	γ, δ	Constant
prob	Probability of individual being overeducated or undereducated	η	Increasing function of human capital
P	Personal characteristics	ε	Error term
Pri	Primary school or below	β	Parameter/coefficient

List of Abbreviations

AAUW	American Association of University Women
ADB	Asian Development Bank
ASEAN	Association of Southeast Asian Nations
AIPEG	The Australia-Indonesia Partnership for Economic Governance
APO	the Asian Productivity Organization
BHPS	The British Household Panel Survey
CHNS	The China Health and Nutrition Survey
Coef	Coefficient
CPS	Current Population Survey
CSAL	Compulsory education program
CSEs	Clustered/Robust Standard Errors
DKI	Daerah Khusus Ibu kota (Capital Province)
DSA	Direct Self-Assessment
EMX	Mixed/ Alternative methods
ESRC	The Economic and Social Research Council
EAs	Enumeration Areas
EU	European Union
FE	Fixed effect
FNSEM	The Fourth National Survey of Ethnic Minorities
Freq	Frequency
GB	Great Britain
GBP	Great British Pound Sterling
GDP	Gross Domestic Product
GED	General Education Development
GER	Gross Enrolment Rate
GNI	Gross National Income
H0	The null hypothesis
H1	The alternative hypothesis
HCT	Human Capital Theory
HH	Household
HILDA	The Household Income and Labour Dynamics in Australia
HSBC	The Hongkong and Shanghai Banking Corporation
IALS	International Adult Literacy Survey
IDR	Indonesian Rupiah (Indonesian currency)
IFLS	The Indonesia Family Life Survey
IIA	Independence of Irrelevant Alternatives
ILO	International Labour Organization
IMR	Inverse Mills Ratio
INPRES	Primary school construction program
IRR	Internal rate of return
ISA	Indirect Self-Assessment
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupation
IV	Instrumental Variables
IV-FE	Instrumental Variables - Fixed Effect
JA	Job Analysis
JICA	Japan International Cooperation Agency
Lao PDR	The Lao People's Democratic Republic
LFS	Labour Force Survey
LM	Lagrange Multiplier

LR	likelihood-ratio
M	Match
MAR	Missing at Random
Max	Maximum
ME	Marginal Effect
Min	Minimum
ML	Maximum Likelihood
MNL	Multinomial Logit
MNP	Multinomial Probit
MoEC	The Ministry of Education and Culture
MoRA	The Ministry of Religion Affair
MP	Marginal Productivity
N/A	Not available
NCDS	National Child Development Survey
NER	Net Enrolment Rate
NGS	National Graduates Surveys
NLC	Negotiating the Life Course
NLS	National Longitudinal Surveys of Labour Market Experience
NLSY	National Longitudinal Survey of Youth
NSSO	National Sample Survey Office
Obs	Observations
OE	Overeducation or years of surplus schooling
OECD	Organisation for Economic Cooperation and Development
OLS	Ordinary Least Squares
ORU	Overeducation–Required–Undereducation
PhD	Doctor of Philosophy
PIAAC	Programme for the International Assessment of Adult Competencies
PISA	Programme for International Student Assessment
PSID	The Panel Study of Income Dynamics
PSLM	Pakistan Social and Living Standards Measurement
PV	Present value
RE	Random effect
REQ	Required years of schooling
RM	Realised Method
SAKERNAS	The National Labour Force Survey (Indonesia)
SSA	Social Security Administration
SCELI	The Social Change and Economic Life Initiative
SD	Standard Deviation
SE	Standard Error
STEP	The World Bank's STEP Skills Measurement Program
SUR	Seemingly Unrelated Regression
SUSENAS	The National Socioeconomic Survey (Indonesia)
UE	Undereducation or years of deficit schooling
UK	United Kingdom
UNESCO	United Nations Educational Scientific Cultural Organization
US	United States
USD	United States Dollar
WEF	World Economic Forum
WERS	Workplace Employment Relations Survey

Chapter 1 General Introduction

1.1 Research Background

The relationship between education and the labour market has been studied extensively by economists over recent decades and across a range of countries, including Indonesia. These studies have been motivated mainly by a search for the causes of wage disparities in the labour market between individuals with different education levels rather than skills.

Indonesia is one of the largest economies in Southeast Asia with a population of around 249 million people; this makes the country the world's fourth most populous. Moreover, the country has a rather young population with around half of the total population below the age of 30 years. Combined, these two features imply that Indonesia currently has a large labour force, around 68 per cent of the total population in 2000 which then increases steadily to 69 per cent or around 125.3 million people in 2014. As such, the human capital development becomes indispensable if the country wants to achieve a sustainable economic growth. Furthermore, one key element of the human capital formations is formal education.

On the supply of labour, the proportion of labour force with higher education qualifications increases. According to the labour force survey data from 2007 and 2017, the proportion of labour force with senior high school qualifications increases significantly from 17.14 per cent in 2007 to 30.84 per cent in 2017. The proportion of workers with undergraduate qualifications also increases from 2.53 per cent in 2007 to 7.48 per cent in 2017. However, the non-schooling population still persists, accounting for around 4.5 per cent of the national population (aged 15 years or older) in 2016 (World Development Indicators, 2016). The Net Enrolment Rate (NER) data of the senior high school and university levels also confirms this increasing trend, from 7.9 per cent in 2000 to 20.18 per cent in 2014 for university level; and from 39 per cent to 59 per cent for senior high school level in the same period.

On the demand of labour, education becomes the most important factor in career progression, particularly in the formal sector; 40 per cent of the job vacancies formally registered with employment in public offices requires minimum qualifications of tertiary education, a further 40 per cent requires job seekers to have completed high school and 20 per cent requires junior high school qualification or lower (Allen, 2016). This is consistent with the economic transformation from agriculture and manufacturing to the service sectors. Thus, there exists a gap between the labour demand and supply characteristics.

Despite the education expansion, one of the education problems in Indonesia is the persistence of low-quality education. Students who take part in the Indonesian Programme for International Student Assessment (PISA)[1] perform below the 25th percentile of the Organisation for Economic Cooperation and Development/OECD's average (World Bank, 2018). Compared to the neighbouring countries, Vietnam and Malaysia are in the 25th -50th percentile; and Singapore leads in the 75th percentile. This low education quality is related to the relationship between the public and private schools in the country. Private schools play an important role in complementing the state-provided education and in helping to meet the demands that the public-school system has been unable to meet, particularly in poor and rural areas, and at the levels of schooling above basic education. However, the quality of the education offered in private schools is generally lower than in public schools, with an exception to elite private schools catering to wealthier families (Tobias *et al.*, 2014). Low education quality and low education level of the population could affect workers performance in the labour market, since they face more difficulties in accessing quality and highly productive jobs. In fact, Indonesia's value of labour productivity is around USD 24.3 in 2015, which is still lower than Thailand (USD 26.5), Malaysia (USD 55.7), and EU15 (USD 81.3)[2].

The government has implemented many policies to improve both school attendance and education quality since the 1970s. The first massive education policy is the 61,000 primary school's construction programmes during 1973-1978, followed by the six-year compulsory education for primary school in 1984. This was then extended by the nine-year compulsory education in 1994 and supported by an abundant allocation for education expenditure from the total government expenditure (20 per cent), which manifests in school grants programs, scholarship system for students from poor families, teacher certification programmes, and revisions of the curriculums.

---

[1] PISA is the OECD's Programme for International Student Assessment. Every three years it tests 15-year-old students from all over the world in reading, Mathematics and science. The tests are designed to gauge how well the students master key subjects in order to be prepared for real-life situations in the adult world (OECD, 2018).

[2] Labour productivity level by individual worker's GDP at constant basic price per worker, using 2011 PPP and reference year of 2015. Data source: APO Productivity Database (2017).

Although these policies have good aims and comprehensive concepts, but the challenges are in the implementations. The school construction programme is relatively successful in increasing the primary education participation at the national level (Duflo, 2001 and Suryadarma and Rogers, 2004). In contrast, the nine-year compulsory education has a slower progress than the six-year compulsory education and the school construction programme (Suryadarma *et al.*, 2006 and Lewis and Nguyen, 2018). Meanwhile, the teacher certification programme may have improved the teachers' living standard as an increase in remuneration is an elemental part of it. However, it does not necessarily increase the teachers' skills or the students' learning performance (Fahmi *et al.*, 2011 and the World Bank, 2018).

These challenges in policy implementations could affect the education sector as well as the labour market outcomes. As a result, the return to education decreases between 1993 and 2007/08 (Purnastuti *et al.*, 2013). Dumauli (2015) argues that one of the possible explanations for the decline in the return to education in Indonesia is the low quality of the education system which affects the quality of the graduates. This could also result in education mismatch in the labour market. In fact, around 3 to 4 out of 10 workers experience a mismatch between their attained education and jobs (Antara, 2017). It is worth noting that education mismatch could also be led by the job and the worker's characteristics (Battu and Sloane, 2002; McGuinness, 2006; and Filsi *et al.,* 2014), but there is still an absence of evidence for this explanation in Indonesia as most studies have focused on the education mismatch trend overtime (International Labour Organization, 2017 and Allen, 2016).

Education mismatch may lead to an increase in unemployment and cause inefficient allocation of the resources invested in education, as well as suboptimal income in individual level or even the existence of penalty wage. Such mismatch may manifest in overeducation or undereducation, either one will be costly to the economy, the firms and the individuals. As a matter of fact, a year increase in the incidence of undereducation among young workers is found to decrease productivity (Kampelmann and Rycx, 2012). Yet, most studies in Indonesia still focus on the effects of overeducation and undereducation on wages for workers with higher education (see *e.g.* Alisjahbana *et al.,* 2017). These studies may not be aware that a prominent issue in the labour market in Indonesia is undereducation of workers with lower education levels, since around 60 per cent of the population has junior high school or lower qualifications. Thus, the present study sets out to analyse the effect of education mismatch on wages from the lowest to the highest education levels, to analyse the return to education with the latest data available; as well as to analyse the determinants of education mismatch. This study also elaborates on the implications of the findings and provides some recommendations to minimise the mismatch from individual level as well as related public policies. Correspondingly, the President of the Republic Indonesia, Joko Widodo asserts that the focus of his second term (2019-2024) will be on the human resource development (Kompas, 2019). Hence, the analysis and evidence on the relationship between the labour market and education qualifications from this study aim to contribute as considerations for policy making in relevant sectors.

1.2 Aims and Research Questions of the Thesis

Research Aims and Research Questions

Based on the research background, the objectives and questions asked in this study are developed following the recent developments in Indonesia. The present study assumes that the education sector expands recently and technology changes rapidly in the country. In the supply and demand framework, these factors will decrease the return to education and increase overeducation. On the other hand, Indonesia still has problems with the non-schooling population and the low quality of education. Thus, the relationships among return to schooling, mismatch incidence and its return are still worth examining. Furthermore, there is still an absence of evidence on the matter in some areas in Indonesia, such as the evidence on the determinants of education mismatch. The waged sectors (both of the public and private sectors) are chosen for the analysis here since these sectors have the same characteristics, namely workers receiving wages regularly. In addition, according to ILO (2000), the basis used for distinguishing formal jobs is that they are covered by the framework of regulation in terms of the enterprise's establishment, company size (commonly medium and large firms) and official registrations. In terms of labour legalisation, the job must subject to the labour law, minimum wage and protected in the job contract between employers and employees.

More specifically, Chapter 2 reviews the education expansion, the education system and its development, as well as discusses the education policies and their implementations.

The chapter also provides data on the labour market development which support the analysis of the next chapters.

Chapter 3 investigates and further updates the return to education in the waged sectors in Indonesia. Specifically, the chapter aims to investigate how the return to education varies across education levels and across time (between 2000 and 2014 periods), to provide a more systematic model specification of the return to education in Indonesia (including religion and ethnicity variables which are rarely observed in previous studies), and to provide an extended analysis by separating the sample by gender and sector. The research questions of Chapter 3 are: (3.1) What are the estimated returns to education in 2000 and 2014? (3.2) Do gender and sector affect the return to education? and (3.3) How does the return to education change between those periods?

Chapter 4 examines the determinants of education mismatch (vertical mismatch[3]) which includes overeducation and undereducation in Indonesia. The aims are to investigate and to update the existing literature on education mismatch, particularly in the waged sectors in Indonesia for the 2000 and 2014 periods; to investigate the variables which affect the mismatch in the waged sectors; and to explore the change on the mismatch determinants between these periods. The research questions of Chapter 4 are: (4.1) Does education mismatch (both undereducation and overeducation) exist in the waged sectors in Indonesia? (4.2) What are the estimates of education mismatch in 2000 and 2014? How does the aggregate trend of education mismatch change between these periods? Are there any distinctions among genders and sectors? (4.3) What are the variables which determine undereducation and/or overeducation? And are there any distinctions among genders and sectors?

Chapter 5 examines the effects of education mismatch on wages. The aims of the chapter are to investigate the extent to which mismatch incidences affect the wages; to see whether overeducated and/or undereducated workers receive lower wages compared to matched workers; to see whether mismatch incidences still affect the wages when unobserved heterogeneity is taken into account; to investigate the effect of gender and sector on returns associated with education mismatch; and to contribute to the existing literature on education mismatch and its returns by taking into account the influence of

[3] A vertical mismatch occurs when the education level of the employee's qualification is not the one required by the job.

unobserved heterogeneity. The research questions of Chapter 5 are: (5.1) Does education mismatch (undereducation and/or overeducation) contribute to determine wage in Indonesia? (5.2) Considering the unobserved heterogeneity, does education mismatch still contribute to determine wage, and (5.3) Do the returns associated with education mismatch differ by gender and by sector?

1.3 Summary of Datasets and Methodology

There are several data sources provided by the government and non-government institutions. The Indonesian government, through the official Statistics Bureau, has some surveys with large-scale sample, such as The National Socioeconomic Survey (SUSENAS) and the National Labour Force Survey (SAKERNAS). There is also a more comprehensive database, *i.e.* The Indonesia Family Life Survey (IFLS). The IFLS is a collaborative effort between RAND Corporations and several institutions in Indonesia. The main difference between the data is provided in Section 2.2.1.

In terms of periods, the survey waves of the 2000 and 2014 are considered when examining the effects of the National Education System Law (No. 20 of 2003) and the Higher Education Law (No. 12 of 2012) enforcement in Chapter 3 and Chapter 4. Thus, the year 2000 represents the period before the law was enforced and 2014-2015 represents the conditions after the law was effective. In Chapter 5, the data from 2007 is added to increase the number of individuals and narrow down the gap between periods in the analysis, thus allowing the present study to analyse a period of around 7 years of changes in wage. Therefore, the conditions before the education expansion period as well as the immediate and longer effects of those policies can be addressed.

In terms of methodologies, Chapter 3 investigates the return to education in the waged sectors in Indonesia by using Mincer wage equations with the Ordinary Least Squares (OLS) method. Mincer wage equation is commonly used to measure wage, particularly in relation with education levels. Mincer equation has been used to estimate wage for different countries, specifications, periods and datasets (see e.g. Psacharopoulos and Patrinos, 2004). The principle of Mincer equation is derived from the Human Capital theory; wage is determined by the worker's educational attainment and experience, excluding training. Meanwhile, the OLS is a simple method to estimate the equations. To analyse the trend overtime, the present study tests the equality of coefficients across two models of Mincer wage equation. Thus, the change can be justified. Admittedly, those methods have some issues, such as endogeneity and sample selection bias. The Instrumental Variable (IV) is used to gauge the role of the omitted variables (particularly the ability bias) or the endogeneity problem in the OLS model. Moreover, Two-Step Heckman models are used as the corrective measures to avoid the possibility of sample selection bias. As such, the results may be considered for the whole target population. While ignoring this correction, it means that the result is valid only for the sub-population of people who decide to work in the labour market. The dependent variable is the natural logarithm of hourly wages, and the main explanatory variables are education level, experience and its squared. The other control variables are relatively more comprehensive than previous empirical studies, which comprises of personal characteristics, work related, firm size, and regional variable categories.

Chapter 4 examines the determinants of education mismatch. Unfortunately, there is no standard model to be used in this area of interest in the literature. Thus, the present study develops a model based on the assignment theory; mismatch is determined by the worker's and job's characteristics. The dependent variable is match categories (1: overeducated, 2: matched; and 3: undereducated). The explanatory variables in the model are determined by the explanatory variables mostly used in previous empirical studies (see *e.g.* Battu and Sloane, 2002): personal characteristics, work related and firm size, and regional dummy variables. In terms of the mismatch determination, the present study applies objective measures, *i.e.* the Realised Method (RM) by calculating the mode and standard deviation.

The Multinomial Logit Model (MNL) is used as a method to analyse the determinants of education mismatch. MNL is also adopted by some studies, such as Kiker *et al.* (1997). The rationale for using MNL is that it allows the analysis of more than two categories of the dependent or outcome variables. The model is also used to predict categorical placement in or the probability of category membership on a dependent variable based on multiple independent variables. Since multinomial logit coefficients can only be interpreted in terms of relative probabilities, the present study also provides marginal effect (ME) calculation to reach conclusions about the actual probabilities. Sensitivity tests of the result are also performed by comparing the mode and mean results, as most studies use mean as a base of their analysis. The present study also carries out the multinomial probit (MNP) model and adds casual workers into the dataset and then compares those results with the results from MNL (main model).

In addition, a hybrid occupation category is developed in this thesis to provide a more accurate measurement of the overeducated, undereducated and matched categories; and to provide a good balance between a strong sample size and reducing the level of heterogeneity in roles within occupational grouping. It is worth noting that the definition of overeducation or undereducation is relative; an individual defined as overeducated in one job may not be so defined in another job. For instance, a worker with 13 years of schooling is overeducated if he is an agricultural and animal husbandry worker (the mean of the worker's education is 8.8 years). However, the worker becomes undereducated if he works as an engineer (which has an average schooling of 14.55 years).

Chapter 5 applies the Overeducation-Required-Undereducation (ORU) model proposed by Duncan and Hoffman (1981) with a panel analysis. The model is an extension of Mincer wage equation with education variables that are decomposed into required years of schooling (REQ), undereducated (UE), and overeducated (OE). Thus, the model enables exploring the relationship among years of required schooling, years of deficit schooling (UE) as well as years of surplus schooling and wage. The model also passes statistical testing in several countries for several periods (see *e.g.* Tsai, 2010 and Dockery and Miller, 2012). There are several advantages from using Duncan and Hoffman's model, *i.e.* the model can provide the estimations necessary to evaluate wage from the monetary perspective. The model also allows the analysis of the return to both undereducation and overeducation. Moreover, the model offers a better interpretation of mismatch and wage relationship; it allows the analysis of economic value to each additional deficit or surplus of education, not just higher/lower wages compared to matched workers.

Furthermore, the panel data model provides a main advantage since it can be used to deal with unobserved heterogeneity. There are two methods in the panel model, *i.e.* fixed (FE) and random effects (RE). The main distinction between FE and RE is whether or not the unobserved individual effect embodies elements that are correlated with the regressors in the model, not whether these effects are stochastic or not. The RE model assumes that the individual-specific effects are not correlated with the independent variables. Meanwhile, the FE model can be used to analyse the impact of the variables that vary over time.

The dependent variable is the natural logarithm of hourly wages, similar to Chapter 3. The main explanatory variables are years of required schooling, years of deficit schooling and years of surplus schooling. The other control variables used are job experience and its squared, personal characteristics, work related and firm size, and regional dummy (similar to previous chapters). And finally, to test the result, the present study employs three approaches as the robustness test, *i.e.* by replacing mode in the main model by mean, by using balanced panel data, and by applying an alternative model from Verdugo and Verdugo (1989).

It is worth noting that the control variables across the chapters may different, Chapter 3 has five religion, six ethnicity, and four firm size dummy variables; Chapter 4 and 5 have only two ethnicity (Java or majority and non-majority ethnicity) and three firm size variables; these changes is conducted considering the finding and its interpretation in previous Chapter; for example: Chapter 4 has some adjustment considering Chapter 3 result. Similarly, Chapter 5 is conducted considers the results in Chapter 3 and Chapter 4, the main aim is simplification of analysis.

1.4 Contribution of the Thesis

This thesis is hoped to contribute to the existing literature by studying the link between educational choices and wage in the Indonesian labour market. The main contributions are:

(1) Chapter 3 provides an analysis of the return to education in the waged sector includes public-waged sector whereas most empirical research in Indonesia only focuses on the return to education in general and on gender disparity. Furthermore, the public sector in Indonesia is particularly important, with 4.5 million civil servants in 2016. The sector also continues to attract job seekers thanks to their appeals, such as: attractive allowances, opportunities to pursue education abroad (Masters and PhD levels), flexible working time, clear career paths, and pension plans;

(2) Chapter 3 also contributes empirically in testing Mincer wage equations in developing countries with a relatively comprehensive set of control variables, consisting of personal and household characteristics, job related and firm size, and residence variables. In terms of endogeneity, besides conventional instruments

such as parents' education, this chapter provides some alternative instruments to deal with ability bias, for instance, policy instruments and health behaviour (smoking);

(3) Chapter 4 provides an analysis of undereducation, both the determinants and their relationship with wage. Undereducation becomes an important issue in the Indonesian labour market since some studies find that the trend increases (see Chapter 4) where more than 60 per cent of the population have junior high school or lower educational qualifications. Undereducation incidence indeed affects the economy, as an increase of one year in the incidence of undereducation can decrease productivity (Duncan and Hoffman, 1981);

(4) Chapter 4 also provides a model of education mismatch determination based on the assignment theory (mismatch is determined by the worker's and job's characteristics) since there is no standard model in the literature. Moreover, the definition of education mismatch is relatively sensitive to occupation digit category. Thus, developing a hybrid occupation category that is specifically adjusted for the research questions gives a better estimation than the major group (1-digit category) or the unit group (4-digit category) of International Standard Classification of Occupation (ISOC) in terms of retaining the homogeneity of sample with large observation and in maintaining the minimum number of observations for the sample with small observation at once. Principally, the hybrid category is rearranged by keeping the sub-categories which have relatively large number of observations and merging the remaining (particularly sub-categories with sample size less than 30 individuals);

(5) the panel data method used in Chapter 5 can contribute to enrich the methods used to analyse wage and education mismatch in Indonesia, since most of the previous studies in Indonesia uses a cross-section model (see *e.g.* Alisjahbana *et al.,* 2017);

(6) and finally, Chapter 5 also contributes empirically in testing the ORU model in developing countries. Considering the sensitive estimation results when using panel data methods, a cautious interpretation is needed in this case. Significant different results between RE and FE may indicate that the ORU model used in this study is relatively sensitive to the panel data methods used. Alternatively, it may point that unobserved heterogeneity may affect the wages.

1.5 Structure of the Thesis

The structure of this thesis is as follows: Chapter 2 is divided into two parts. The first part details the recent developments in Indonesia; the education expansion, the education system and relevant policies on education; as well as the labour market developments. The second part concerns the data, which provides background of the data selection (IFLS), as well as its advantages and disadvantages. Next, Chapter 3 discusses the return to education in the waged sectors in Indonesia by using Mincer wage equation. The other important aspect of analysis is trend overtime, which is also detailed in the chapter. This chapter also provides some robustness test results which address the endogeneity issue and the sample selection bias. Turning to Chapter 4, the determinants of education mismatch are presented and analysed. The analysis is conducted based on undereducation and overeducation. Furthermore, to make the results more reliable, some sensitivity tests using other methods (mean and multinomial probit model/MNP) are performed and casual workers are added into the dataset. Chapter 5 further elaborates on the effects of education mismatch on wage by using panel data. To provide an extended analysis, Chapter 3, 4 and 5 perform the models by separating the sample by gender and sector. Finally, in the last part of this thesis, Chapter 6 offers the conclusions, implications as well as recommendations of this research.

Chapter 2 Developing Countries and Indonesia Context

Indonesia is one of the largest countries in Southeast Asia, having the largest archipelago in the world which consists of 17,508 islands; five of them are the major islands. The country also has a population estimated at 270.63 million in 2019, increases around 5.3 per cent from the 2015 estimate of 257 million. This also makes Indonesia the 4th most populous country in the world, after China, India and the United States (World Bank, 2019). Currently, the country is considered a middle-income country. With abundant human resources, human capital development becomes one of the most important factors to develop the country.

This chapter comprises two parts; the first part is on the country's context, which discusses some issues in the education system, the education expansion, demographics and the labour market developments. Those issues set the background and give empirical data for the study on the return to education, educational mismatch and its relationship with wage. The second part is on the data, which discusses several alternative data that are available to be used in analysing the labour market and the education system in Indonesia. Moreover, this part also provides some background of the datasets chosen (the IFLS) and their advantages as well as disadvantages for the present study.

2.1 Developing Countries and Indonesia's Context

2.1.1 The theories of Labour Market in Developing Countries

The theory that commonly used to analyse in the labour market in developing countries is from Wachter (1974) and Cain (1976). The theory stressed that for labour market dualism to exist, different wages must be paid in different sector to comparable workers.

Lewis (1954) in Fields (2005) reveals earnings in the subsistence sector set a floor to wages in the capitalist sector, but in practice wages have to be higher than this, and there is usually a gap of 30 per cent or more between capitalist wages and subsistence earnings. The gap is "illusory" because of the higher cost of living in the capitalist sector, there remained a real wage gap due to 1) the "psychological cost of transferring from the easy going way of life of the subsistence sector to the more regimented and urbanized environment of the capitalist sector," 2) the payoff to experience in the capitalist sector, and 3) “workers in the capitalist sector acquiring tastes and a social prestige which have conventionally to be recognized by higher real wages.”

Moreover, Kuznets (1955) in Fields (2005) further developed the model of wage dualism and intersectoral shifts by exploring how various measures of income inequality would change as the high-income sector comes to employ an increasing share of the population. All of the inequality measures used by Kuznets exhibited an inverted-U pattern, which later came to be known as the “Kuznets Curve.” In additions, Lewis and Kuznets should not be faulted for neglecting the human capital issue, because human capital theory had not yet been devised when they developed their dualistic development models.

The more modern labour market dualism literature stressed that for dualism to exist, different wages must be paid in different sectors to comparable workers (Schultz (1961, 1962), Becker (1962, 1964), and Mincer (1962, 1974)). These theories will be elaborated in literature review part (Section 3.2). Since the late 1960’s and early 1970’s, a whole class of models has arisen in which a wide variety of wages exist in the labour market, and workers are presumed to search among employers for the best possible opportunities. There are some critics on the dualistic labour market model, such as: lack of a precise behavioural interpretation of the results that is the principal shortcoming of the dualistic labour market empirical studies (Rosenzweig, 1988); the formal and informal economies operated in segmented labour markets and there is limited mobility between the two (IADB, 2003). Nevertheless, labour market dualism was a most useful starting point for analysing some economies when it was first presented decades ago, and it remains a useful characterization of some economies today.

In the development, the segmented labour market approach divides the markets into three components: the formal sector labour market, the informal sector labour market, and the links between the two. The first model is the formal sector labour market, Fields (2005) presents four alternative models of wages and employment in the formal economy: a) the market-clearing labour market model, the standard labour market model; b) models with wages set above market-clearing levels for institutional reasons, or model with a minimum wage; c) models with wages set above market-clearing levels for efficiency wage reasons; and d) models with wages set above market-clearing levels because of supply-side considerations. The other models that identified are matching models (Mortensen and Pissarides, 1999), job creation and destruction models (Davis, Haltiwanger, and Schuh, 1996), ranking models (Blanchard and Diamond, 1994); and imperfect information models (Stiglitz, 2002).

Secondly, informal sector labour market, this is based on the fact that people are working informally because they are unable to work formally. Fields (2005) divides the informal sector labour market into: the informal economy as a desirable sector and the informal economy with its own internal dualism. The previous empirical research reveals that in Mexico and other Latin American countries, self-employment in the informal sector may be more attractive for a majority of workers compared to work as wage employees in the formal sector (Maloney, 2003, 2004; World Bank, 2007).

And the last component is intersectoral linkages in the labour market. Fields (2011) identified some models in this component: (1) a noncompeting groups model in which individuals belong to one labour market segment or another, and they cannot or will not switch from one to another (as in many human capital models). (2) An integrated labour market model which starts with two or more sectors but assumes that all of the equilibrating forces that apply to a single labour market with market-clearing also apply to a labour market with a multiplicity of sectors, so that wages equalize across sectors; nearly all international trade models have such a labour market specification (Krugman and Obstfeld, 2003). (3) The crowding model, which assumes that any worker who is not employed in the high wage part of the economy takes up employment in the low wage part of the economy; the Lewis and Kuznets models mentioned above can be recognized as crowding models. (4) One sector maintains higher wages than another, but the two sectors are linked via workers' job search behaviour, such that in equilibrium the expected wages associated with the two sectors are equal to one another (Harris and Todaro, 1970). (5) Workers choose occupations which maximize their current and future returns, but because of imperfect capital markets, occupations that require high levels of investment cannot be entered by persons with low initial wealth.

2.1.2 Indonesia's Context: Education Definition and Institutions

Education Definition

Based on Law No. 20/2003 on the National Education System, Indonesia's national education system consists of formal, non-formal (*e.g.* soft skill courses) and informal educational systems (*e.g.* lifelong learning). Formal education includes several types of education such as general, religious, vocational, professional, and special education. It comprises several levels, starting from early childhood education, primary school-6 years education (with the average age of 7-12 years), junior secondary schools-3 years education (with the average age of 13-15 years), senior secondary school-3 years education (with the average age of 16-18 years), and higher education.

Education Institutions

There are two types of junior and senior high schools in the country: general and vocational. General senior high schools are aimed at those who intend to go on to university, whereas vocational junior/senior high schools are for those who plan to find jobs right away. Islamic institutions could also provide those education levels, but the difference is those institutions have more Islamic subjects and content in the curriculum, as well as provide general subjects such as science and social science modules.

For vocational high school (SMK), there are 13,710 vocational high schools in 2017/18, with 75 per cent of them being managed by the private sector. The tuition fee of SMKs is higher than general senior high schools (SMA) since most are managed by private sector and they have more laboratories/practice activities (70 per cent of learning activities) which require much more cost than SMAs. In terms of location, most SMKs are located on Java Island, the most populated island. In terms of their study programmes, most of these SMKs specialise in engineering and technology, business and management as well as information and communication technology/ICT (SMKs Statistics, Ministry of Education, 2017/18).

For the higher education level, there are five types of institutions, *i.e*. academy, polytechnic, college, institute, and university. The first two specialise in vocational education streams, whilst the last three are more comprehensive and allowed to offer all education streams. A college (Sekolah Tinggi) is a specialised institution focusing on one particular academic discipline. Unlike universities, institutes are specialised in a particular group of disciplines such as sciences and technologies, arts, or agriculture.

Universities always provide education at the level of *Sarjana I* (Sarjana Stratum Satu, S1), equivalent to an undergraduate degree. These programmes have a nominal length of 4 years. Students are allowed to extend their study duration to a maximum of 14 semesters or 7 years). The *Sarjana II* (Sarjana Stratum Dua, S2) is equivalent to a master degree and requires a minimum of 2 years/39-50 credit points following an undergraduate degree, including a research and 8-10 credit points for a final paper. Finally, the PhD programme requires 40-52 credits with a nominal duration of 2-2½ years (4-5 semesters) for candidates holding a master's degree. For candidates holding a bachelor's degree, the required amount of credits is 76-88 with a nominal duration of 4-4½ years (8-9 semesters). Generally, students finish a PhD programme in 3-4 years, following master degree (Nuffic, 2017).

Apart from those three core degrees, universities also offer diploma degrees (higher professional education programmes) which require less time than the 4 years programme, for instance, diploma 1 requires only one year of higher professional education, diploma 2 two years, and diploma 3 three years of education. There is also diploma 4 which is comparable to the bachelor's degrees in a similar specialisation. Moreover, most Indonesians do not continue straight away from undergraduate to higher-level educations; they usually work for several years and return to university for post-graduate programmes after possessing working experiences.

There are several criteria for a public university entry: one must have already graduated from a senior high school in the last three years and passed the university entrance exams. There are several entrance exams that can be chosen by students such as the public university entrance joint selection (usually at the national level), independent selection (at the university level) and invitation based (at the university level). The first two selections are based on written test. The last selection is based on the grade score, national examination score, or other academic achievements. Meanwhile, private university entrance process is simpler; only based on their discretion and selection mechanism.

In additions, based on the Indonesia Report on Standards and Qualifications (2015), there is a national qualifications framework under presidential decree in 2012, the Indonesian Qualification Framework (IQF), that provides nine qualification levels with each level based in four main dimensions: (1) job skills, (2) science or knowledge comprehension, (3) capacity to select and apply methods and knowledge and (4) management skills. The framework has a strong focus on equivalency and recognizing all pathways to obtaining qualifications and promotes recognition of prior learning.

Indonesian National Certification Agency (BNSP) is established as as a certification authority in charge of implementing the certification of personnel and professional competence for labour. Some of their functions are: to give license to LSP by process of accreditation to ensure and maintain performance of the certification bodies; to ensure implementation of the Qualification Framework certification scheme is in place for assessment and certification; to approve the development of the certification scheme, especially for clusters and competency units; and to control and monitor certification implementation.

2.1.3 Education System and Expansion

Education System

In Indonesia, two ministries are jointly in charge of managing the education system, *i.e.* the Ministry of Education and Culture (MoEC) which is responsible for over 80 per cent of all students, teachers and schools, and the Ministry of Religion Affair (MoRA) which is responsible for the remaining 20 per cent comprising Islamic schools or madrasah (Statistics Indonesia, 2019).

In terms of institutions, the Indonesian education system has diverse providers, with private schools estimated to serve 31 per cent of all students and to employ 38 per cent of all teachers. Private education also comprises over 30 per cent of upper-secondary enrolment and over 80 per cent of tertiary enrolment (Statistics Indonesia, 2019). In Indonesia, these schools play an important role in complementing the state-run education and help to meet demands that the public-school system has been unable to meet, particularly in poor and rural areas and at levels of schooling above basic education. However, the quality of education offered in these schools is generally lower than in public schools, with the exception of elite private schools catering to wealthier families (Tobias *et al.*, 2014).

Hendajany (2016) compares the effectiveness between public and private schools using the national exit exam of Junior Secondary Schools in Indonesia. Using the Indonesian Family Live Survey (IFLS) in 1997, 2000 and 2007, this study finds evidence proving that graduates of the public school have higher scores on the national exit exam than those of the private school, whilst controlling a wide variety of students' characteristics and family background.

Through Ministerial Decree No. 51/2018 regarding the school enrolment zoning system, a substantial improvement in school’s quality could be achieves since the regulation ensures a proportional student distribution across schools and thus putting an end to parents' tendency to favour particular schools. The policy is effective for both public and private schools, with the exemption of vocational schools due to the limited number of schools and majors.

Education Policies and Its Empirical Studies

Indonesia has had several government policies that may have had a significant effect to increase school participation from primary school to university level. The policies for primary schools are the school construction program in 1973; the six-year compulsory education for primary school in 1984; and the nine-year compulsory education in 1994 (as an extension of the previous program). A 12-year compulsory education program (the Universal Secondary Education) has been planned since 2013, but these plans have not yet been implemented due to the associated costs and other reasons (WENR, 2019).

The first policy that has a significant effect on education development in Indonesia is the primary school construction program, known as the Sekolah Dasar INPRES. The government launched this program in 1973, based on the Presidential Instruction 10/1973 with the aim of achieving the equality of primary education across provinces in the country. Throughout this program, more than 61,000 primary schools were constructed between 1973/74 and 1978/79. This affected the increase of school-per-pupil ratio; an average of two schools per 1,000 children aged 5 to 14 in 1971. This program is reported as the fastest primary school construction program ever undertaken in the World (Duflo, 2002).

Empirically, Duflo (2002) also points out that the school construction program leads to an increase in education among individuals who were young enough to attend primary school after 1974, but not among the older cohorts. The study uses the number of primary schools built between 1974 and 1978 as the instrument and finds an increase of 10 percentage points in the proportion of primary school graduates in the labour force, which increases their formal labour force participation by 4–7 per cent. Suryadarma *et al.* (2004) assert that the government implemented a 6-year Compulsory Education program for the first time in 1984, which proved highly successful; by 1988, 99.6 per cent of children were either enrolled in primary schools or had finished the six-year compulsory education.

The other significant policy is the six-year compulsory education for primary-school-aged children (the 7-12 years old population), based on the Presidential Instruction Decree No. 10 of 1973, which was fully implemented in 1984. The term “compulsory” suggests that a six-year education should be universal and that every Indonesian child should have the right to at least six years of education. As a result, the government was responsible for providing adequate educational facilities and went on to build many schools, including in remote areas, provided many primary school teachers for these schools. Afterwards, the government launched the extension of the six-year compulsory education which became the nine-year compulsory education later on. This includes primary and junior high schools (the 7-15 years old population).

Some empirical studies find that this program increases Indonesian children’s basic education attainment in terms of primary school and junior high school attendance, given primary school completion. Suharti (2013) notes that the gap in education attainment between children from poor and rich households becomes non-existent at the primary level and continues to narrow down at the secondary level. Thus, those programmes successfully increase the school attendance. The program may also affect the decrease in the prevalence of child labour in Indonesia. But, the up-to-date data are not available and the latest data are from 2009: of the total number of Indonesian children aged 5-17 years, namely 58.8 million, it is estimated that 6.9 per cent of them (around 4.05 million) are working children and of this figure, 43.3 per cent (around 1.76 million) are child labours (Statistics Indonesia and ILO, 2010).

Compulsory education also improves the quality of education in terms of literacy rates, pupil-teacher ratio and international test result (Tobias *et al.*, 2014). The younger generation (individuals aged 30 years) also has the highest rates of high school completion and more of them have completed tertiary education. Purnastuti *et al.* (2013) assert that the enrolment of each education level has expanded. The participation rate in primary schools increased from 79 per cent in 1973 to 92 per cent in 1993. In line with that, enrolments at the higher levels of education have also expanded, both as a flow-on effect of compulsory basic education and as a result of direct policy initiatives. Meanwhile, Suryadarma *et al.* (2006) argue that the net enrolment rate at the junior secondary level in 2004 was 65 per cent, only 10 percentage points higher than a decade earlier; the progress is relatively slower than primary schools, in spite of the 9-year Compulsory Education program in 1994. This is due to many factors, such as household welfare, sex, religion, and abundant employment opportunities which all lead to low secondary school enrolment.

Lewis and Nguyen (2018) add that the main reasons for the 9-year compulsory education's failure in Indonesia are that the government did not support the initiative with sufficient additional funding and that it was lax in enforcing the policy. Despite the education expansion, the main problem of low-quality education persists. Akresh *et al.* (2018) analyse the long term and intergenerational effects of school construction programme in 1973 and find that parents exposed to the school construction program transmit these benefits to the next generation. The effect does not occur during primary school, since primary school is almost universal when second-generation individuals attend school. Yet, the effects extend throughout the secondary and tertiary education. Relative to the baseline levels, the largest impacts are seen in tertiary education with effect sizes indicating a 20 per cent to 25 per cent increase in the likelihood of the second-generation child completing university.

Besides, regulations related to higher education levels such as the Education Law (No. 20 of 2003) and the Higher Education Law (No. 12 of 2012) have some significant effects on higher education institutions. In particular, it gives greater autonomy over curriculum, as well as their management and use of resources. For instance, universities can set their own remuneration system based on merit and performance. In addition, the government allows foreign entities to invest in the country's education institutions, albeit in conjunction with local institutions. As a result, the autonomous universities started to enjoy greater freedom to develop their management systems and to manage their sources more efficiently. This law was issued to accommodate the educational demands of tens of millions of Indonesians approaching the university age over the coming years, a formidable challenge that the government is struggling to tackle on their own. The effects of higher education reform are reflected in academic areas such as higher average Grade Point Average (GPA) and faster study completion during 2002-2009 periods (Nizam and Nurdin, 2014). For instance, an undergraduate degree which normally takes around 8-10 semesters (4-5 years) to complete has recently been shortened to 6-8 semester (3-4 years).

Moreover, the Law number 22 of 1999 (later amended by the Law number 32 of 2004, and number 23 of 2014) on regional autonomy have implications on the different authority levels of schools' management. The management of public primary and junior secondary schools in Indonesia is under the responsibility of the district/city administration. Public general and vocational senior high schools are managed by the province administration. The Ministry of National Education only acts as a regulator in the field of national education standards.

Teachers

One of the most important changes is that through the government regulation No. 74/2008, a teacher is defined as a professional educator whose main duties are to educate, teach, guide, train, access, and evaluate students through formal education level, from primary education to higher level of education. There are some requirements for teachers: (1) they must have teaching certification, adequate physical and mental health, and also the capability to achieve the national education goals; (2) they must have academic qualification of at least an undergraduate degree (diploma IV or Sarjana) with compatible subject study (*i.e.* primary school education for primary school teachers); and (3) they must be competent in pedagogy, personality, social and professional education. Thus, primary school teachers who do not have primary school education background must pursue another undergraduate degree or a master degree in primary school education. The regulation is later revised in 2018, where some requirements such as compatible academic qualification and teaching certification are removed (Kompas, 2018).

Many researchers have explored the relationship between teacher certification and the education quality; most of them agree that the certification program may have improved teacher's living standard as a remuneration increase is an elemental part of it. However, it does not necessarily improve the teacher's skills or the student's learning performance[4] (Fahmi, *et al.*, 2011 and World Bank, 2018).

There are two types of teachers in Indonesia: in-service[5] and pre-service teachers. Pre-service teachers or temporary teachers work in private and public schools and are paid by the foundations. These teachers vary widely in their qualifications. In terms of pay structure, Indonesia has two-tier labour market for teacher. The first tier is the in-service teachers; those with an undergraduate qualification and already certified earn around IDR

[4] Political pressures also watered down the certification process and left only the pay increase in place.

[5] In-service teachers comprise several categories: (1) Public teachers are civil servants who have minimum teaching qualifications. (2) Contract teachers are fixed-term teachers who are usually employed through (donor-funded) projects and have the same qualifications as public teachers. And, (3) permanent teachers are engaged by the foundations to teach in private schools. These teachers' qualifications vary by the quality of the school.

7 million per month (GBP 378)[6]. The second tier is pre-service teachers, the salary is far lower than that of the in-service teacher, starting from IDR 300,000 (GBP 16) monthly in rural areas. The total budget for teacher's allowance is more than IDR 60 trillion in 2018 or around 2.7 per cent of total national spending (Cabinet Secretary, 2018).

Another issue is the uneven distribution of teachers although nationally there was an oversupply of teachers; approximately 55 percent of schools have an oversupply of teachers, particularly in most urban areas; 34 per cent of schools are undersupplied in remote and rural areas. Moreover, teachers in rural and remote areas tend to have less education than their counterparts in urban areas (The World Bank, 2008).

Curriculum and Teaching Style

With regards to curriculum, the 2013 curriculum has been gradually implemented since 2014. It interprets a competency-based curriculum which covers attitudes as well as thinking, social and cultural skills. The most important aspect is that the curriculum advocates student-to-student active learning. The curriculum is sufficient and good, even compared to other ASEAN countries; yet, the challenge is in the implementation and the pedagogic style.

In primary to senior high schools, there are many styles of teaching, three of them are (1) the conventional way of teaching which has become a culture within schools, teachers dominate the class by explaining the subject, with teachers continuing to practice this style of teaching from generation to generation; (2) Active, Innovative, Creative, Effective and Fun Learning; and (3) Easy, Fun and Enjoyable Learning. By applying the active learning method, teaching can be perceived as an activity that involves facilitating learning and assisting students to develop their own understandings (Law and Miura, 2015). However, in the implementation, most teachers apply a curriculum with the conventional way of teaching. World Bank (2018) confirms that 60 per cent of the time a typical Mathematics class is spent on lecturing, with limited time remaining for practical

[6] (Assumed IDR 1 = GBP 18484) Comprises basic salary (the same for all civil servants) which is around IDR 2.5 million; professional allowance around IDR 327.000 (GBP 17.7); the certificate allowance of around IDR 95 per cent of basic salary; and other allowances (regional performance allowance) which vary depending on regions, for example: DKI Jakarta is around IDR 7 million (GBP 378.7), West Java is around IDR 2 million (GBP 108.2) in 2018.

work or problem-solving exercises. Subsequently, the improvement in education quality is not optimal.

Education Quality

There are several international indicators that are commonly used, such as the human capital rank from the World Economic Forum (WEF), PISA and TIMSS scores. According to The Human Capital Report 2015 (WEF, 2015), Indonesia ranks 69th out of 124 countries in terms of human capital development. The report quantifies how countries are developing and deploying their human capital and tracks progress over time based on education level, skills and employment availability. Compared to its regional peers, Indonesia is still left behind Singapore (24th), Malaysia (52nd), and Thailand (57th).

Meanwhile, PISA, a measurement by OECD which is commonly used for cross country comparison, shows that the education quality in Indonesia is relatively low. Its performance is in the position below the 25th percentile of the OECD average, as shown in Figure 2.1 (World Bank, 2018). Furthermore, based on Trends in International Mathematics and Science Study (TIMSS)[7], Indonesia's score in international mathematics achievement decreases from 411 in 2003 to 397 in 2015. The top achievers in 2015 are Singapore (618), Hong Kong (615), South Korea (608), Chinese Taipei (597), and Japan (593), while England score is 546 points.

[7] TIMSS is an international assessment of mathematics and science at the fourth and eighth grades, conducted every four years since 1995. In 2015, 57 countries and 7 benchmarking entities (regional jurisdictions of countries such as states or provinces) participated in TIMSS. In total, more than 580,000 students participated in TIMSS 2015. The majority of TIMSS items assess students' applying and reasoning skills.

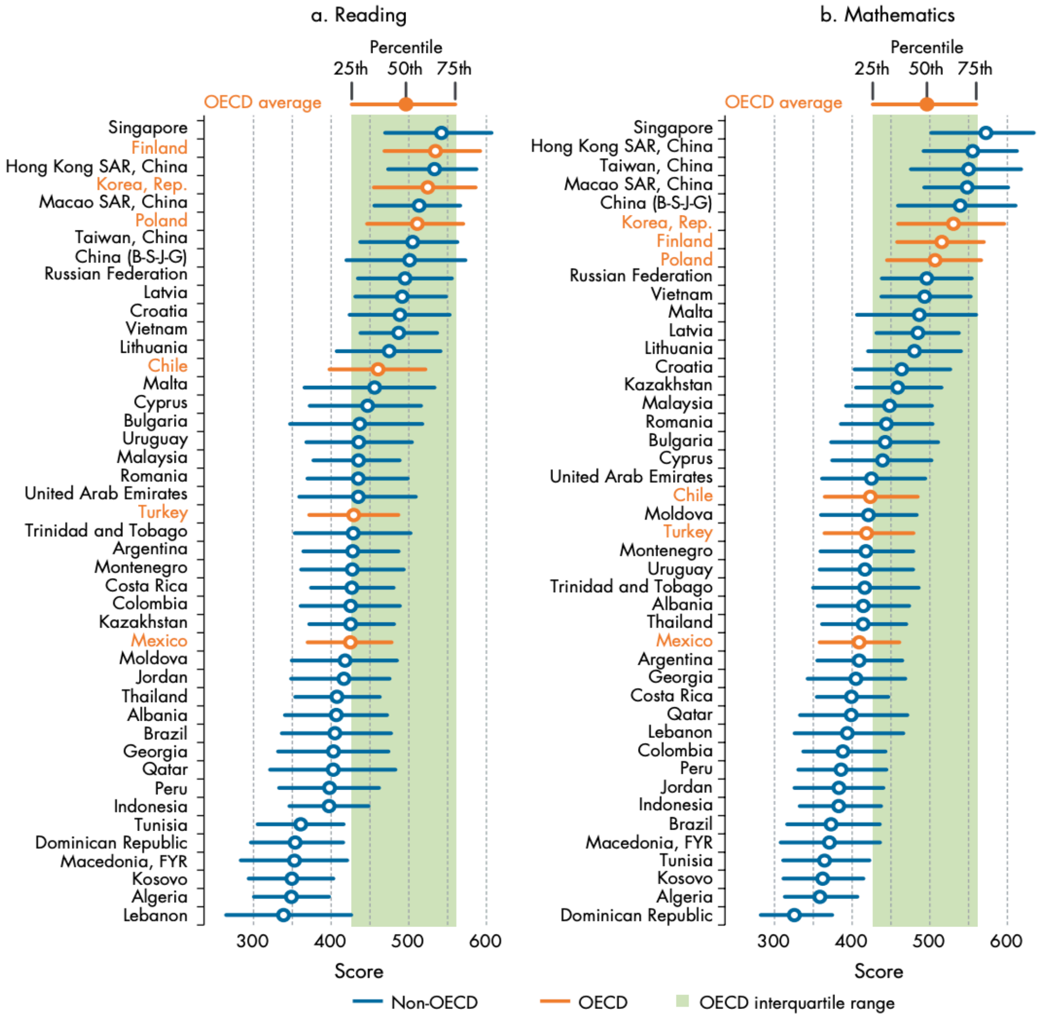


Figure 2.1: PISA Score for Reading and Mathematics, 2015
Source: World Bank, 2018.

Education Expansion

Education has been expanding in Indonesia since 1990s, following the success of school construction and compulsory education policies, which is reflected by the Gross Enrolment Rate (GER)[8], as one of education development indicators (Figure 2.2). Between 2000 and 2014, the highest increase of GER is at university level (151 per cent), from 10.26 per cent in 2000 to 25.76 per cent in 2014. This is followed by senior high

[8] GER is defined as the number of students enrolled in a given level of education, regardless of age, expressed as a percentage of the official school-age population corresponding to the same level of education. For the tertiary level, the population used is the 5-year age group starting from the official secondary school graduation age. Alternative measures of GER is NER (Net Enrolment Rate) which defined as the total number of students in the theoretical age group for a given level of education enrolled in that level, expressed as a percentage of the total population in that age group (UNESCO Institute of Statistics, 2019).

school, from 50.22 per cent to 73.95 per cent in the same period or around 47.3 per cent increase. Correspondingly, the GER of primary school and junior high school in 2014 is already relatively high; 109.2 per cent and 89.98 per cent, respectively. Some possible reasons for the increase of higher education are the changes in demographic structure with decreasing population in the 19-23 years brackets, positively affecting the enrolment rate; there is also a recent increase in the number of institutions in higher education (Moeliodihardjo, 2013). Even though those programs relatively succeed, non-schooling population still persists; around 4.5 per cent of the population (aged 15 years or older) in 2016 (World Bank, 2016).

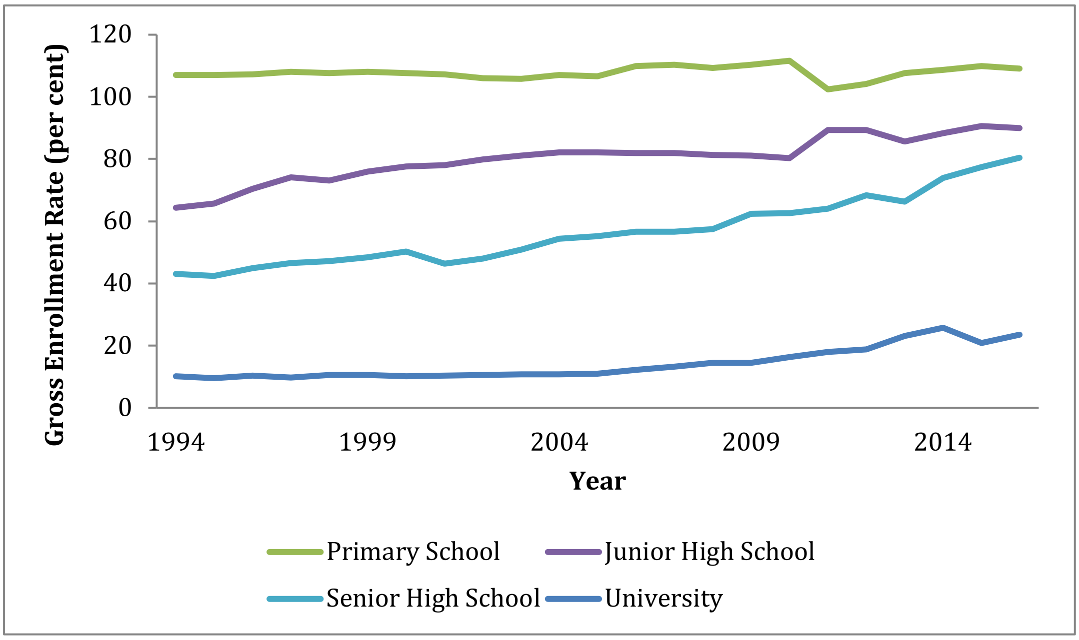


Figure 2.2: Education Indicator: Gross Enrolment Rate, 1994-2016
Source: Statistics Indonesia.
Notes: Gross Enrolment Rate (GER) is defined as the number of students enrolled in a given level of education, regardless of age, for example: GER primary school = number of primary school students divide by population aged 7-12-year-old*100.
The percentage could higher than 100 percent since there is a deviation, primary school student can enrol at age 6 (nearly 7) or student age slightly higher than 12-year-old could also still counted in primary school for several reason such as resit or delay.

Apart from the increase in GER, Gropello and Sakerllariou (2010) confirm that there is a dramatic increase of education attainment in Indonesia between 1994 and 2007, in particular at senior high school and university levels, as shown in Table 2.1. The proportion of workers with university qualifications increases by more than 190 per cent; the proportion of workers with lower secondary and upper secondary general also increases by 24.6 per cent and 59.3 per cent respectively. In contrast, the proportion of workers with primary school or below decreases by 32.7 per cent. Similar trend also occurs in other Asian countries, such as the Philippines.

Table 2.1: The Change in Education Qualification Over Time in Several Asian Countries

Education Level	Indonesia	The Philippines	Thailand	China
	1994-07	1998-06	1990-04	1999-05
Below primary school	-69.8	-33.7	-32.9	-59.3
Primary school	-32.7	-30.9	31.4	-45.0
Lower secondary	24.6	5.3	45.2	-16.6
Upper secondary general	59.3	45.0	100.0	19.3
Upper secondary vocational	-9.6		14.7	
Tertiary	190.3	37.1	98.4	15.2

Source: Gropello and Sakerllariou (2010).

The GER of senior high school and university levels steadily increases, possibly due to some government interventions and some other possible reasons from the demand side, such as the changes in demographic structure with decreasing population in the 15-19 years brackets, which positively affects the enrolment rate, and the recent increase in the number of institutions of higher education (Moeliodihardjo, 2013).

Moreover, Indonesia has more than half of the population under the age of 30. The middle class counts at least 52 million people (20 per cent of the total population) whose consumption accounts for 43 per cent of the total household consumption. Expanding Indonesia’s middle class can help boost the economic growth and broaden prosperity through one of the most important aspects, *i.e.* education (The World Bank, 2017). Moreover, Indonesians have a strong family-oriented culture; that means that householders are more likely to invest in education which is seen as a means of boosting living standards, so much so that consumer spending on this item is higher than in other regional countries. Thus, spending on education increases significantly with income level – the richest households spend more than 25 per cent more on education than the poorest households, but all households allocate around 4 per cent of per capita spending for education (British Council, 2017).

Comparison with other countries

Compared to similar countries, the trend of expansion in education is relatively similar, as shown in Figure 2.3. Primary schools have a very high gross enrolment ratio, around

100 per cent in average. The average enrolment in secondary education level is around 80 per cent in 2015, with Thailand being the highest (120 per cent) and India the lowest (73 per cent). For university or tertiary level, Indonesia has a relatively lower enrolment ratio (23.3 per cent), similar to India (26.8 per cent), whereas Malaysia and Thailand have higher ratio, 42.3 per cent and 45.9 per cent, respectively. It seems that the trend will continue increasing in the future.

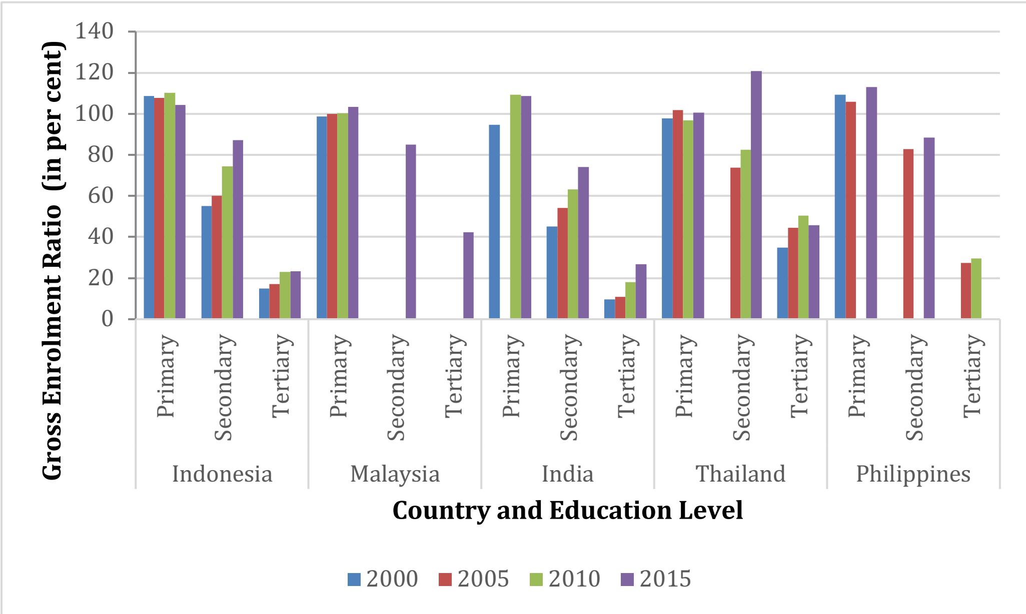


Figure 2.3: Gross Enrolment Ratio in Selected Countries, 2000-2015 (in per cent)
Source: World Development Indicators, 2017.

In terms of pupil-teacher ratio, primary school in Indonesia has a ratio of 16.5 in 2014; this is far lower than in 2000, which implies that there is an improvement. The ratio for lower secondary school is slightly lower than primary school (14.7 in 2014). In contrast, the ratio for upper secondary school is higher than the others, 16.6 in 2014; and the trend increases. One of the possible explanations is that the number of senior high school is far lower than the number of primary schools, and it has not yet been part of the compulsory education programme. In general, pupil-teacher ratios in Indonesia for all education levels are slightly higher than the ratio of OECD; compared to the UK, those ratios are similar in 2014, as shown in Figure 2.4.

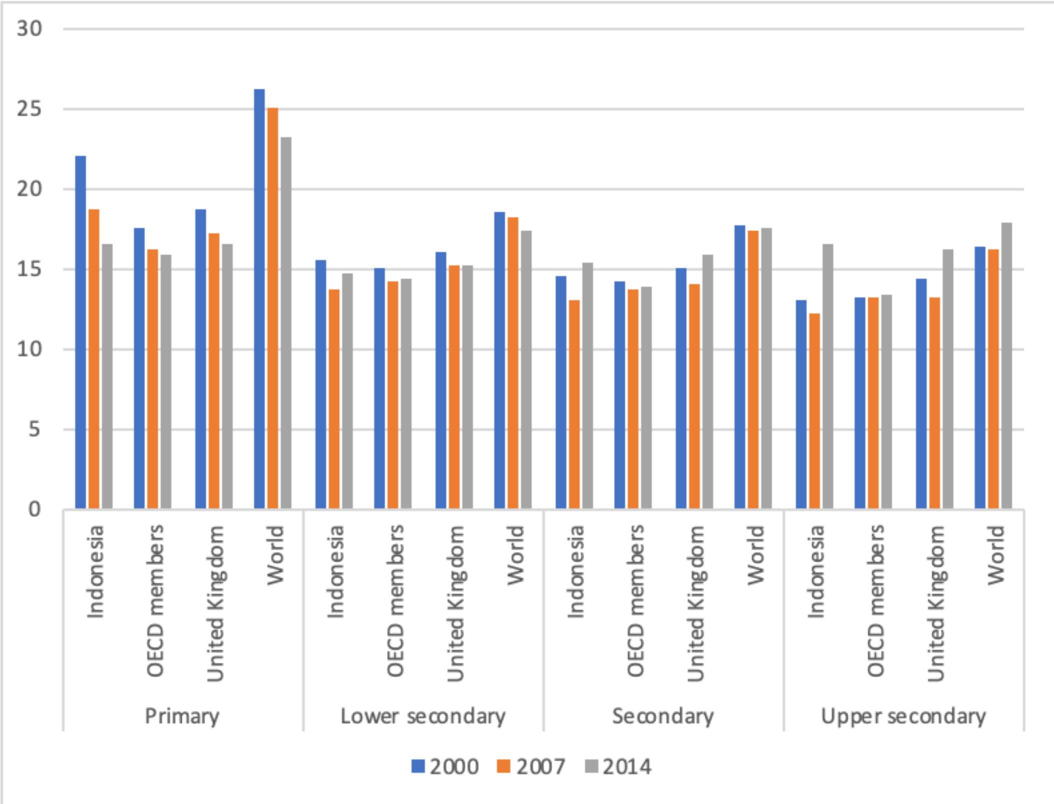


Figure 2.4: Pupil-Teacher Ratio, Indonesia and OECD, 2000 to 2014
Source: World Development Indicators, 2019.
Note: World data is on average.

Despite the education expansion in Indonesia, the population with junior high school or lower qualifications is still far higher than the population with senior high school or higher qualifications (Figure 2.5). Allen (2016) asserts that the composition of the labour force continues to be dominated by workers with lower levels of education. In particular, it is more common for the population aged over 40 years to have primary school or below as their highest level of education. Those aged below 40 years are more likely to have completed junior high school or senior high school. Those aged below 30 years (younger generation) have the highest rates of high school completion and many have even completed tertiary education. Considering that Indonesia is one of the most populous countries and the highest demographic dividend occurrence, human resource plays an important role in sustaining the country's future economic growth and supporting the economic transformation. More investment in human resource is required since highly educated and well-trained human resources are the engine of the economy.

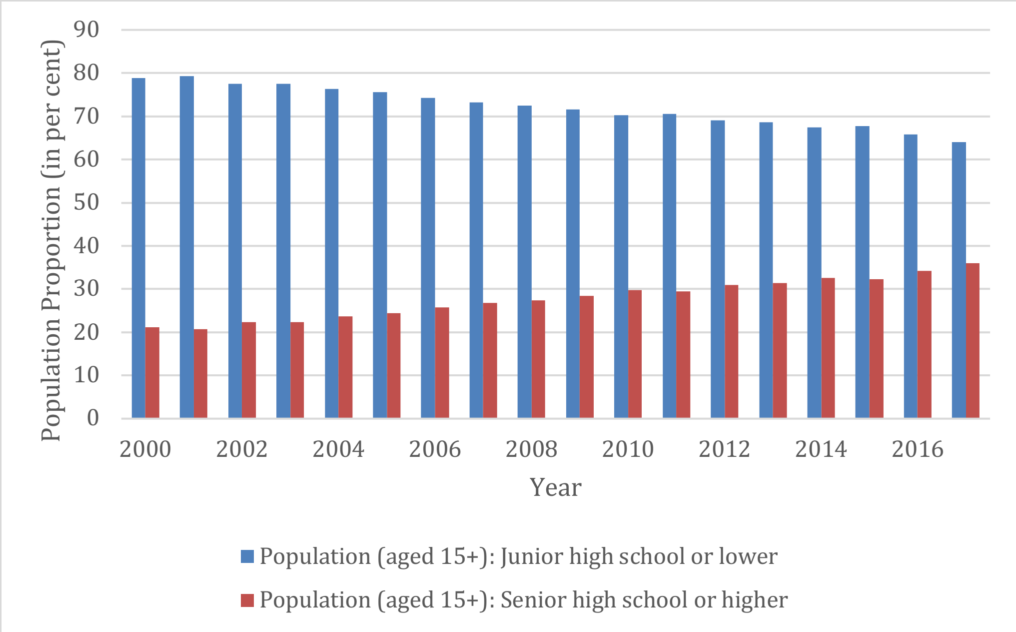


Figure 2.5: Population by Education, 2000-2017
Source: Central Bureau of Statistics Indonesia, 2019

Education Spending

Based on education level, public school's fee for primary and junior high schools is free, due to the nine-year compulsory education. For private schools, the education fees vary according to the location, institution type, and ownership (WENR, 2019). Meanwhile, public senior high schools and universities have a substantial lower education fees and are more affordable than private ones. Spending on senior secondary and higher education has increased between 2008 and 2009 by 26 and 53 percent respectively, but households seem to have absorbed a large portion of the cost of expansion, and total household spending on education in general has increased significantly over the decade (The World Bank, 2013). According to Statistics Indonesia data, average spending per capita on education by the household was 3.6 per cent of total spending in 2000, and increased to 3.9 per cent in 2011, or around IDR 24,679 (GBP 1.4).

In terms of country comparison, HSBC (2017) conducted a survey on the value of education, with the aim to provides authoritative insights into parents' attitudes and behaviour towards their children's education around the world. Over 500 parents (including at least 150 with a child at university or college) were surveyed in all countries. Compared to other countries, Indonesian parents spend around USD 18,422 in total towards their child's higher education (including school/university tuition fees, educational books, transport, and accommodation). This amount was relatively lower compared to parents in the UK, while Hong Kong has the highest spending in education, as depicted in Figure 2.6. It is worth noting that the survey conducted to the parents with child at university or college, that possibly only represents top quantile income group and sample selection bias could occur. While, most of Indonesian education attainment was primary and junior high school, as shown in Figure 2.5. Although the figure represents the high-income group, Indonesian expenditure on expenditure was still lower.

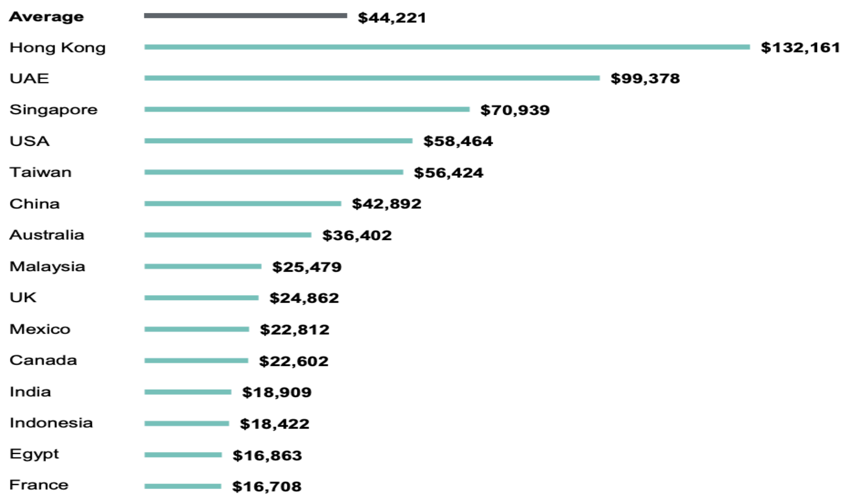


Figure 2.6: Parents' Spending on Their Child's Education, Selected Countries
Source: HSBC, 2017.
Notes: Based on question 'Have you ever paid for private tuition for your child?', base: parents with a child in primary, secondary or tertiary education.

2.1.4 Demographics and Labour Market Developments

Demographics

From the total population of 257 million people, Indonesia is expected to reap from its huge working age group (Figure 2.7), which will reach 70 percent of the total population by 2030. Recently, Indonesia is entering the initial stage of the much-vaunted demographic dividend, which is expected to peak within 12 years (The Jakarta Post, 2018).

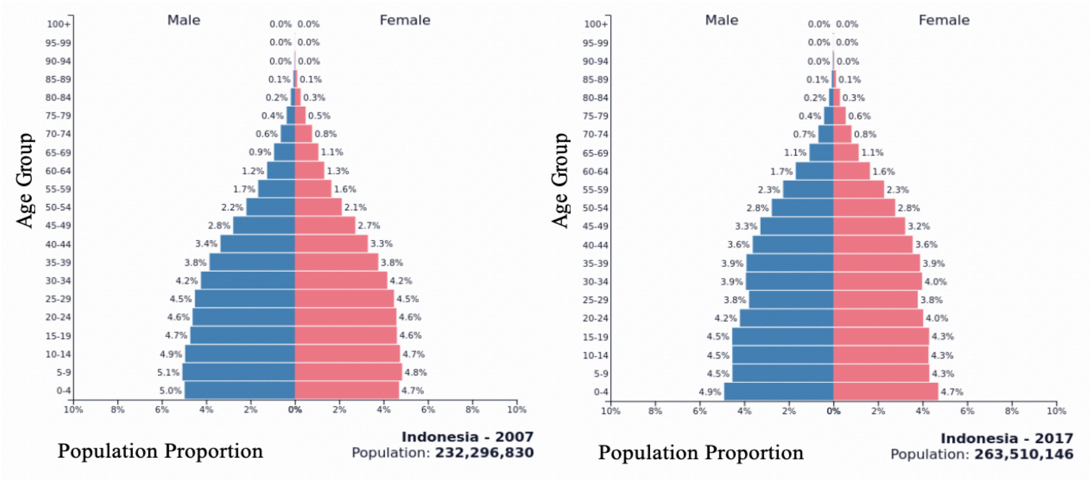


Figure 2.7: Population Pyramid of 2007 and 2017
Source: PopulationPyramid.net

Labour Force

ILO (2014) records show that the Indonesian labour force has reached 127.6 million (around 50 per cent of the population). 46.4 per cent of those employed were working in the formal economy and 53.6 per cent were working in informal employment. Although only 46.4 per cent of employment works in the formal sector, its contribution to GDP is around 60-70 per cent in 2010.

Employment

Statistics Indonesia (2019) records the average unemployment rate being 7 per cent during the 2000 – 2018 periods with a minimum of 5.13 per cent in 2018 and a maximum of 11.2 per cent in 2005. Based on education level, the highest unemployment rate occurs to senior high school (8.29 per cent) and vocational high school (11.41 per cent) graduates, with the share of 27.6 per cent and 24.7 per cent in 2018, respectively, as shown in Figure 2.8. ILO (2014) confirms that unemployment is an issue for the youth, as the unemployment rate for people aged between 15 and 24 years estimated at 17.1 per cent in February 2014. Indeed, the youth account for over 50 per cent of the unemployed population and most unemployed youth have never worked before.

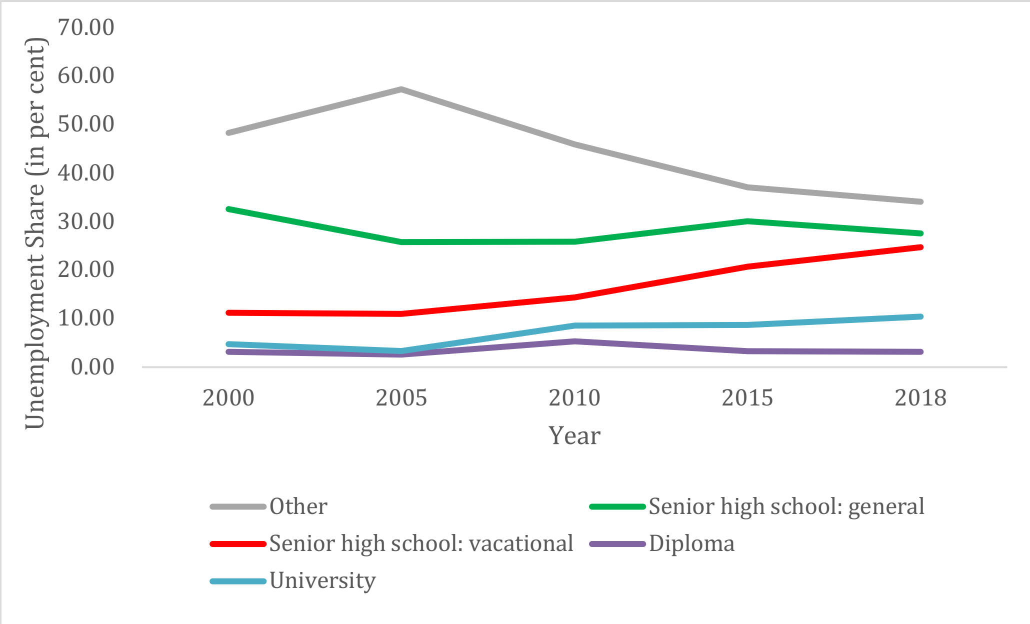


Figure 2.8: Unemployment and Education Levels

Source: Statistics Indonesia, 2019.

Table 2.2: Labour Market Indicators, Indonesia and OECD, 2000-2017

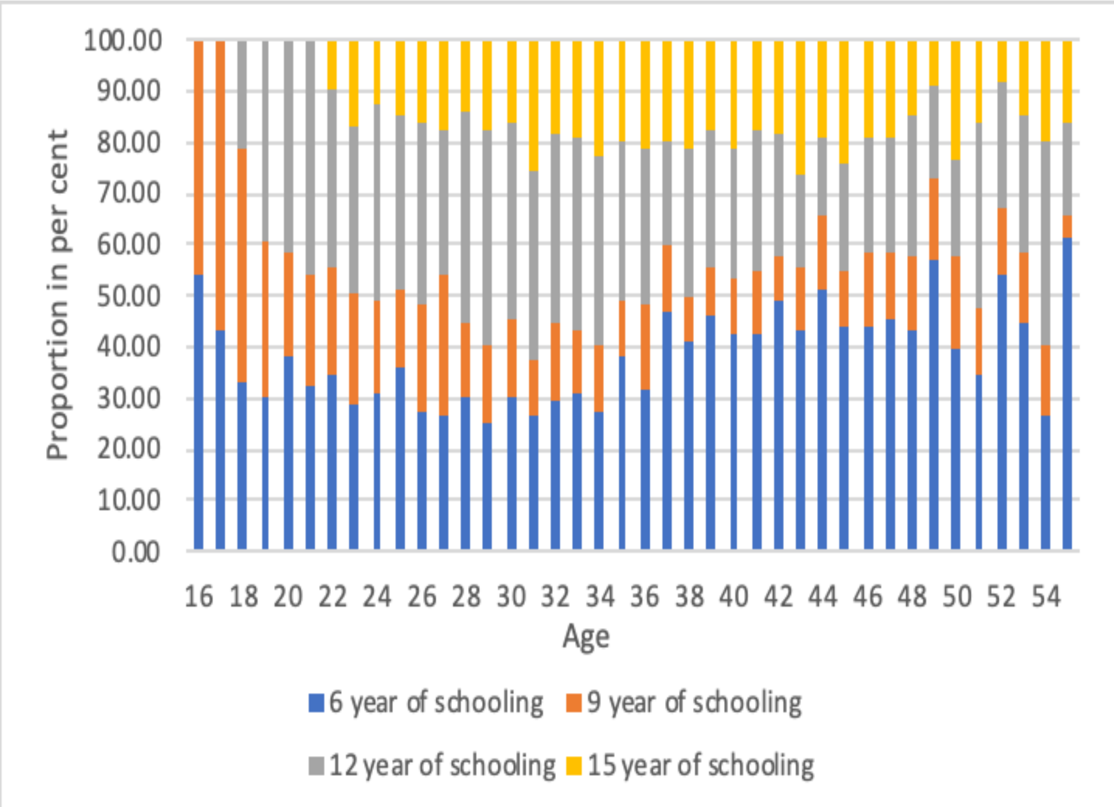
Country Name	Series Name	2000	2007	2014	2017
Indonesia	Employment to population ratio, 15+, total (%) (modelled ILO estimate)	63.2	60.8	64.2	64.4
Indonesia	Unemployment, total (% of total labour force) (modelled ILO estimate)	6.1	8.1	4.0	4.2
Indonesia	Unemployment with intermediate education (% of total labour force with intermediate education)	13.7	17.1	7.6	7.7
Indonesia	Unemployment with basic education (% of total labour force with basic education)	5.1	6.0	2.9	2.9
Indonesia	Unemployment with advanced education (% of total labour force with advanced education)	10.4	13.2	5.0	4.6
Indonesia	Labour force participation rate for ages 15-24, total (%) (modelled ILO estimate)	52.8	50.4	47.2	47.8
OECD members	Employment to population ratio, 15+, total (%) (modelled ILO estimate)	56.3	56.7	55.3	56.6
OECD members	Unemployment, total (% of total labour force) (modelled ILO estimate)	6.3	5.6	7.4	5.8
OECD members	Unemployment with intermediate education (% of total labour force with intermediate education)	9.2	7.0	11.6	9.1
OECD members	Unemployment with basic education (% of total labour force with basic education)	13.0	10.5	15.6	12.4
OECD members	Unemployment with advanced education (% of total labour force with advanced education)	4.6	4.0	7.0	5.9
OECD members	Labour force participation rate for ages 15-24, total (%) (modelled ILO estimate)	51.2	48.7	46.0	46.5

Source: World Bank, 2019.

Compared to OECD, the employment-to-population ratio in Indonesia is relatively high, at 64.4 per cent in 2017. Meanwhile the ratio in OECD is 56.6 per cent for the same period. Unemployment rate in Indonesia is slightly lower; 4.2 per cent of total labour force in 2017. Meanwhile, the rate in OECD is 5.8 per cent. In terms of the labour force participation rate, Indonesia placed slightly higher (of 47.8 per cent) than OECD (46.5 per cent). For unemployment by education, Indonesia has a substantially lower unemployment with basic education qualification (2.9 per cent in 2017), while OECD scores 12.4 per cent in that aspect (Table 2.2).

Turning to the distribution of workers based on years of schooling and age, it seems the younger generations had higher education attainment than the older generation in 2000, the highest education attainment is 15 years of schooling or undergraduate degree (Figure 2.9). Based on gender, young male (22-32 years old) tends to have higher education attainment than female at the same age (Figure 2.10).

Figure 2.9: The Distribution of Workers Based on Years of Schooling and Age in 2000



Source: The author's calculation based on IFLS3 of 2000.

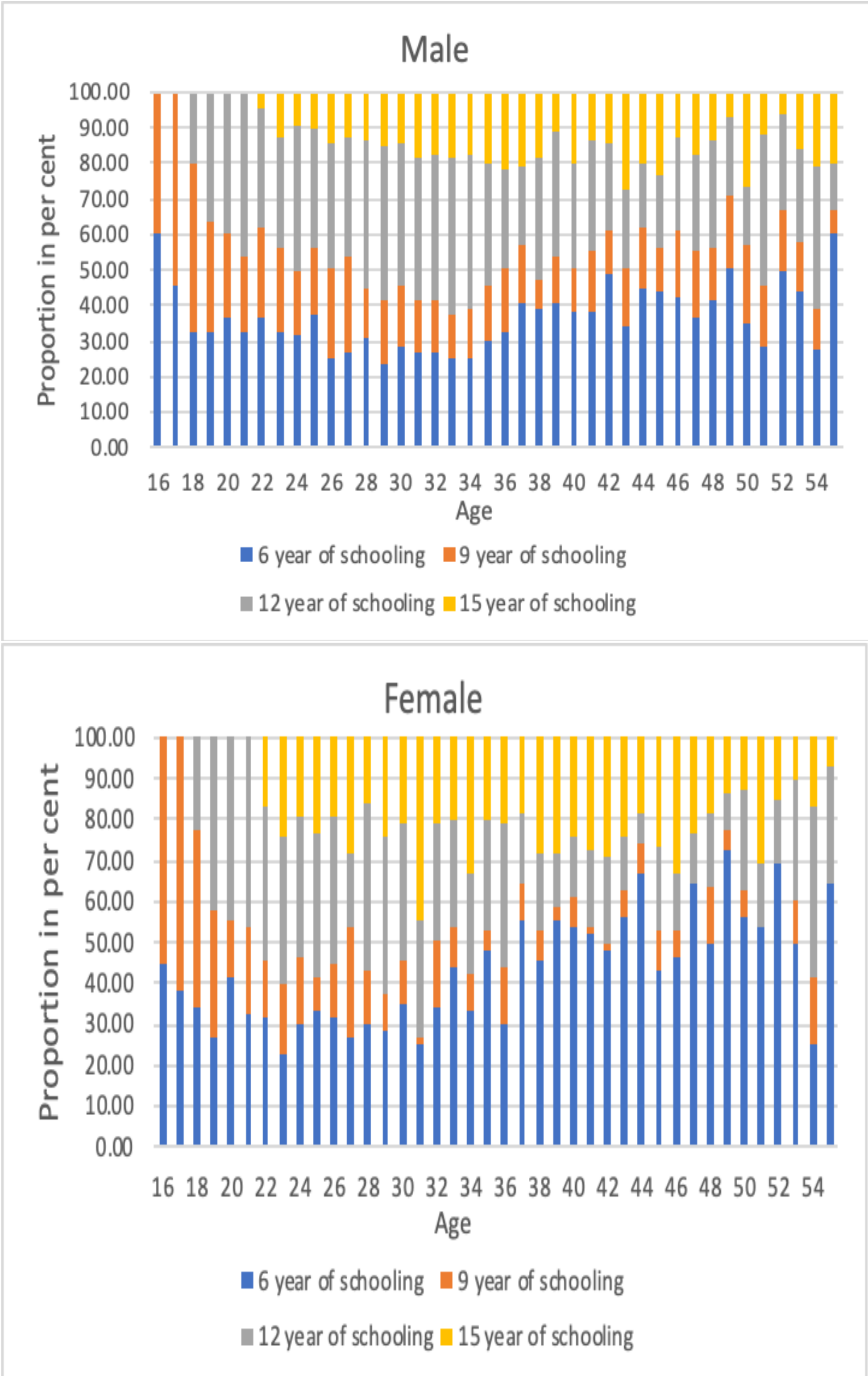


Figure 2.10: The Distribution of Workers Based on Years of Schooling, Age and Gender in 2000
Source: The author’s calculation based on IFLS3 of 2000.

In 2014/15 (15 years later), the highest education attainment increases to 22 years of schooling or PhD level, although the number of PhD is still lower in the population. There is also a slight shift in the distribution, the proportion of workers with 6 and 9 years of school decreased, this also may confirm that education expands in Indonesia (Figure 2.11). Similar trends occur between the gender, the peak of distribution shift to the middle, and male tends to have higher education attainment than female (Figure 2.12).

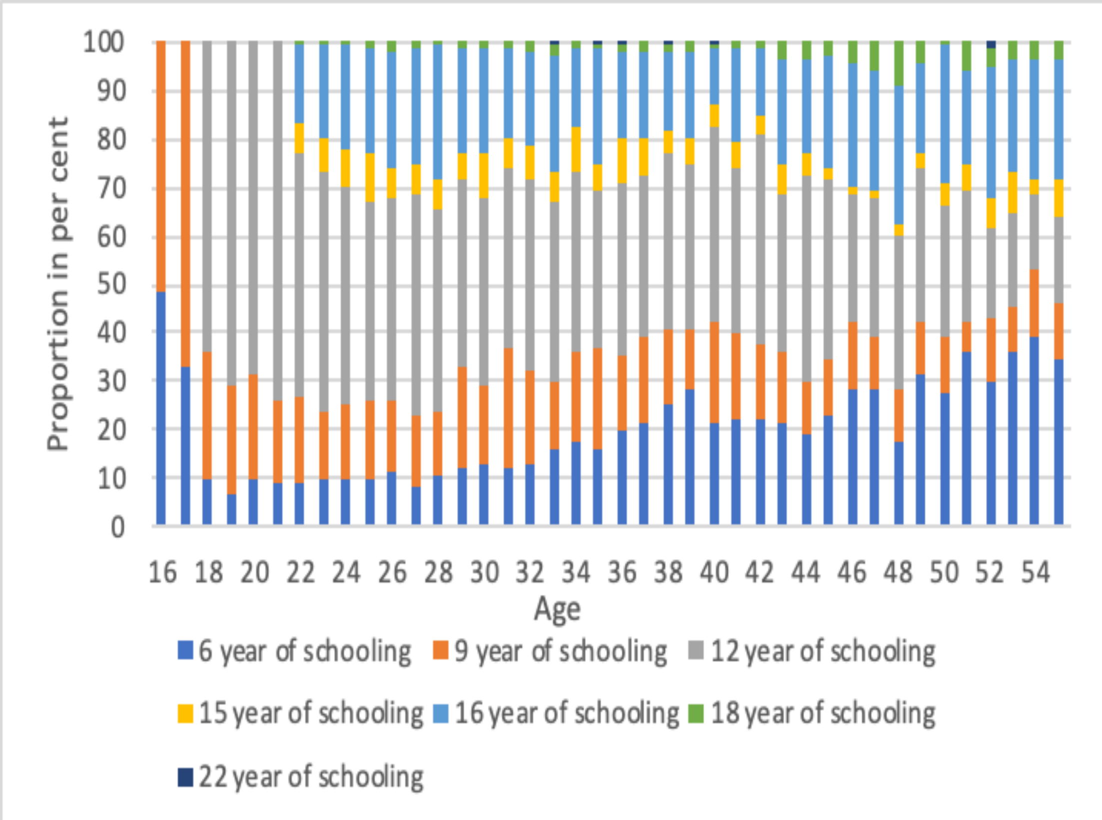


Figure 2.11: The Distribution of Workers Based on Years of Schooling and Age in 2014/15

Source: The author's calculation based on IFLS5 of 2014/15.

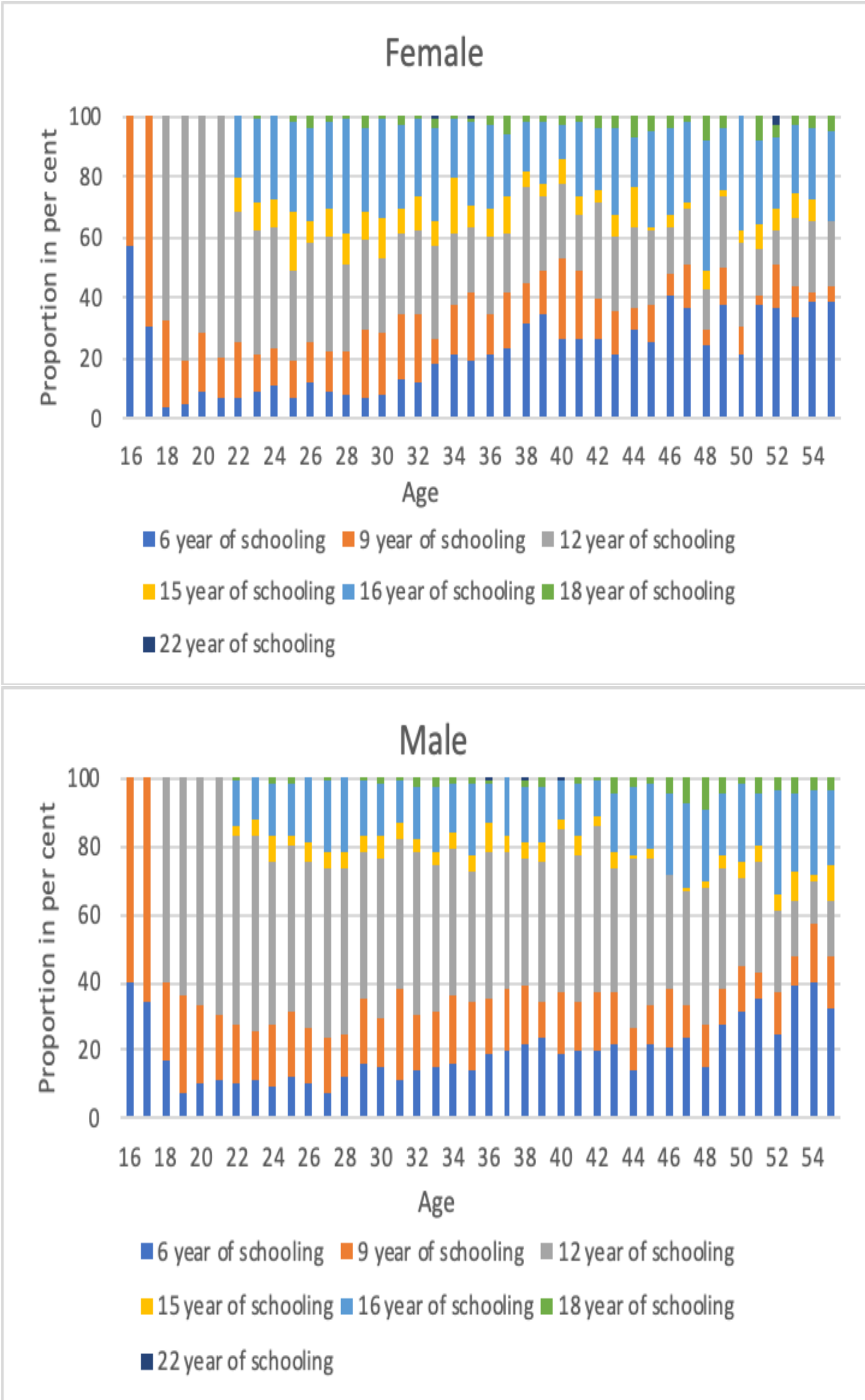


Figure 2.12: The Distribution of Workers Based on Years of Schooling, Age and Gender in 2014/15
Source: The author's calculation based on IFLS5 of 2014/15.

In terms of underemployment, most study in Indonesia focus on work less than normal working hours and would willing to accept another job such as: Dhanani (2004). Moreover, there is a study review educated (university) underemployment in Indonesia (Nagib and Ngadi, 2008). Nagib and Ngadi found the number of educated underemployments reached 3.43 million people (around 12.24 per cent of the total underemployment rate) in 1997. The quantity and percentage of this unemployment type have been rapidly growing from year to year. During 1997- 1999, the number of educated underemployments went up from 3.4 million (12.24 per cent) to 4.27 million people (13.63 per cent) in 1999. The 1997 economic crisis also brought impact on the increased number and percentage of educated underemployment although it was found minor for youth workers. Yet during 2002- 2003, the number of educated underemployments decreased from 3.7 million people (12.66 per cent) to 3.62 million people (12.72 per cent) in 2003; and increased to 3.87 million people (13.85 per cent) in 2004. Thus, it seems the timing of labour market entry (e.g. individuals who enter the labour market just after the Asian financial crisis) affect educated underemployment in Indonesia.

Economic Transformation

In terms of sector, job creation changes in line with economic transformation from agriculture to the service sector. Jobs in agriculture fell to 34 per cent in 2015 from 56 per cent of all employment in 1990, while works in the service sector have surged to 53 per cent from 34 per cent, and manufacturing jobs have also increased from 10 per cent to 13 per cent. In the past decade, the service sector grows extraordinarily, creating around 22.2 million out of 23.8 million (93 per cent of the total job creation) new jobs between 2001 and 2015. The manufacturing industry also gained 3.2 million jobs (13.4 per cent), in contrast to agriculture and mining which lost 1.6 million jobs or -6.7 per cent (Wihardja, 2016).

Formal and Waged Sectors

According to the Indonesian Ministry of Labour and Immigration (2011), the number of job seekers who registered with employment offices (in terms of stock) was 1,941,434 individuals (23 per cent of estimated unemployment) as of 2011, as shown in Figure 2.13. This is significantly smaller than the number of individuals who were estimated to be unemployed by Statistics Indonesia in 2011, which comprises 8,319,779 individuals (ILO, 2014), meanwhile 1,094,729 vacancies were registered in the ministry. In general, the supply of labour outstrips the demand for it. In contrast, the demand for labour outstrips the supply of labour for university level. As such, the issue of education mismatch is a high possibility in Indonesia.

In terms of occupations, the majority of employment has junior high school (9 years schooling) or lower qualifications, and most of them work in agricultural industries, production and related or sales. Meanwhile, most of university graduates work as professional, technical and related workers, as well as clerical and related workers (Allen, 2016).

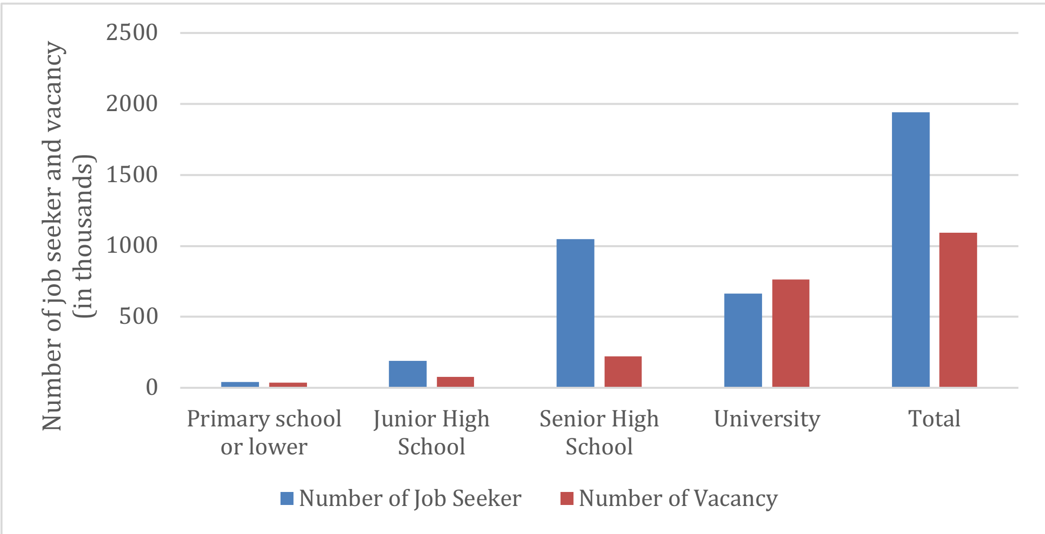


Figure 2.13: Registered Job Seekers and Number of Vacancies in Indonesia in 2011
Source: Dit. PKK Ditjen Binapenta, The Ministry of Labour and Immigration, 2011.

Public Sector

Government or public sector plays an important role in the economic development through allocation, distribution and stabilisation, as well as its role in the labour market. The government has estimated 4.5 million civil servants in 2016 or 1.7 percent of the total population (The Jakarta Post, 2016).

In terms of wages, Indonesia’s salary structure is similar to an egalitarian system in the past, resulting in most of its best graduates from well-known and highly qualified universities uninterested in becoming government employees. Moreover, the low salary tends to encourage wrongdoings and illegal activities such as accepting bribes and asking for compensation for services provided (Tjiptoherijanto, 2014). In 2013, in line with the issuance of Law No. 5/2014 on State Civil Apparatus, public organisations such as bureaucracy have undergone reform to be structurally lean and functionally rich. As a result, public sector workers are offered the same wages nationwide, and the wages in public sector are as competitive as in the private sector. In addition, there are certain occupations in Indonesia where being a public sector worker provides a substantial salary boost such as teachers and lecturers.

Furthermore, the public sector continues to attract job seekers due to their appeals, such as attractive allowances, opportunities to pursue education abroad (Master and PhD levels), clear career paths and flexible working time; consequently, there will be more opportunities to be with family or have an additional job and earn more money for a living (Tjahjono, 2017). Meanwhile, Saputra (2018) argues that being a public sector worker can lift one's social status to the highest level, bringing more respect from society. Thus, the civil service is still regarded by many Indonesians as a far better job than any other occupations even though working for a bank, for instance, could be more lucrative. Another reason is the associated privileges. In Indonesia, public sector workers enjoy easier access to various services such as banking, mortgages, and health care. It has been the rule of the game in Indonesia that banks or other financial institutions are much more willing to lend money or give credit to those workers because civil servants have a magic letter called a "decree." This letter guarantees financial stability and security, for the simple reason that civil servants are paid by the country, not by a corporation.

Regarding the recruitment process in the public sector, LaForge (2016) asserts that before the reformation era in Indonesia (1997/98), entry into the civil service was based on a single examination, after which an employee was essentially guaranteed a job until retirement. In the past, *promotions were based* on the close relationship and personal loyalty of an employee to their superiors (Sakinah, 2017), or (*simply) willingness to pay* as observed by Blunt (2012) in Sakinah (2017), that positions in the civil service were seen as 'tradable goods' open to the highest bidder; *rather than merit or ability or the worker's performance. Moreover, regulations made it nearly impossible to fire a civil servant, except in cases of criminal doings. Even so, an arduous appeals process made it difficult to remove employees who broke the law*. In 2014, the government stipulates Law number 5/2014 on civil servant, several significant changes are related to merit-based recruitment and open promotion. However, the main challenge is in the implementation. KPPOD (2017) in KSI (2017) added even within the internal bureaucracy, some bureaucrats are not happy with the implementation of this law, especially the meritocracy and open promotion system for Senior Executive Services.

Remuneration was another issue; disparities between the private and public wages drive corruption in the public sector. Thus, the civil service law was amended in 1999 to stipulate that the remuneration of public employees should be comparable to that of private employees. Following bureaucracy reform in Indonesia, the government through law number 5/2014 states procurement of civil servants are through planning, vacancy announcement, applying, selection, result announcement, probation, and finally the appointment. Each government agency openly announces to the public their need to fill positions with prospective civil servants. Every Indonesian citizen has an equal opportunity to apply for a civil servant position after fulfilling the requirements. The selection is processed by the government agencies through an objective assessment based on competence, qualification and other requirements as governed by the position's characteristics. The reform also means that the recruitment process should be objective, transparent, and accountable. According to the State Civil Service Body's data, the number of applicants in 2017 is around 2.5 million people, while the vacancies available are only around 37.000. Thus, the probability of being accepted is around 1:65.

Wages

Based on GNI per capita (constant 2010 USD) in 2014, Indonesia has USD 3572.6, the highest one is Norway with USD 92147.4, world's GNI per capital is USD 10152.4, meanwhile OECD's GNI per capita is USD 37722.1 (World Development Indicators, 2019).

In the private sector, Amiti (2011) affirms that wages in Indonesia are largely determined by the market, with the exception of minimum wages set by provincial governments. Meanwhile, the determination of public service's wage is more complicated, since the public service's wage scale is organised according to seniority, position, rank, and political approach; consisting of a base wage in addition to several layers of allowances, and even personal relationship. Some of these allowances cover family, food and housing needs; others relate to the type of position held, and still others serve as rewards to top management in exchange for loyalty (World Bank, 2000). In terms of wage comparison in public and private sector, it seems public sector has premium wage since 1990s. While, the lowest wages in public sector is slightly lower than the average wage in private sector, as shown in Figure 2.14.

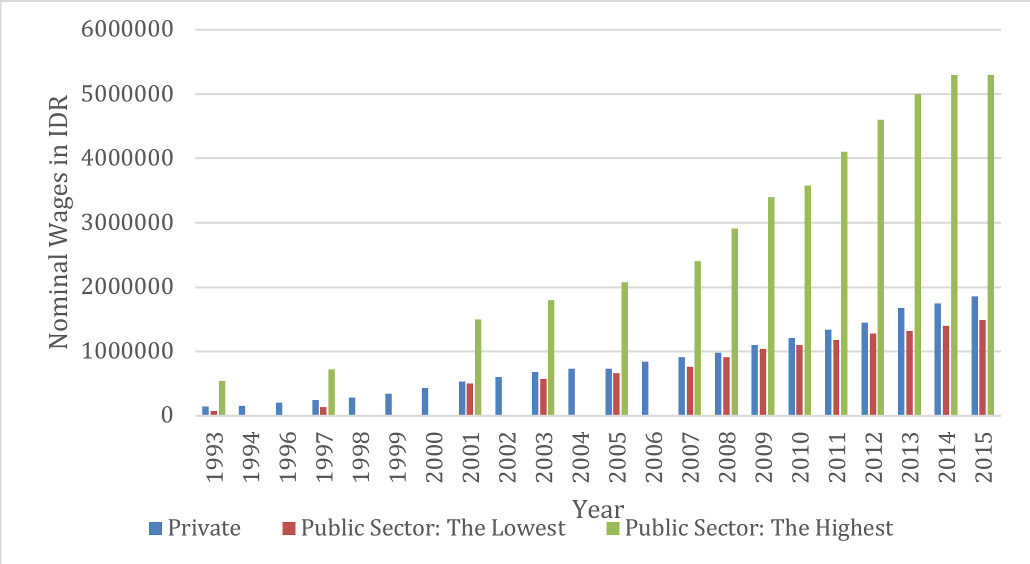


Figure 2.14: Wage Comparison in Public and Private Sector, 1993-2015
Source: CEIC Database and Government Regulation on Public Sector Wages, Various Years

Turning to minimum wage, the determination of minimum wage is based on Article 1 (1) of the Minister of Manpower and Transmigration Regulation No. 7/2013 which states that the minimum wage is the lowest wage consisting of basic wage including fixed allowance set by the governor as a safety net. SMERU (2002) reveals that minimum wages are binding for the bulk of workers in the formal sector. Minimum wage in Indonesia varies; the lowest is in Yogyakarta (Central Java Province), around USD 102.7 per month; and the highest is in DKI Jakarta (capital province), around USD 258 per month. Compared to other ASEAN countries, the lowest minimum wage in Indonesia is relatively competitive, almost similar to Myanmar; but only a few provinces have a low minimum wage. Meanwhile, Malaysia has relatively high minimum wage (Figure 2.15).

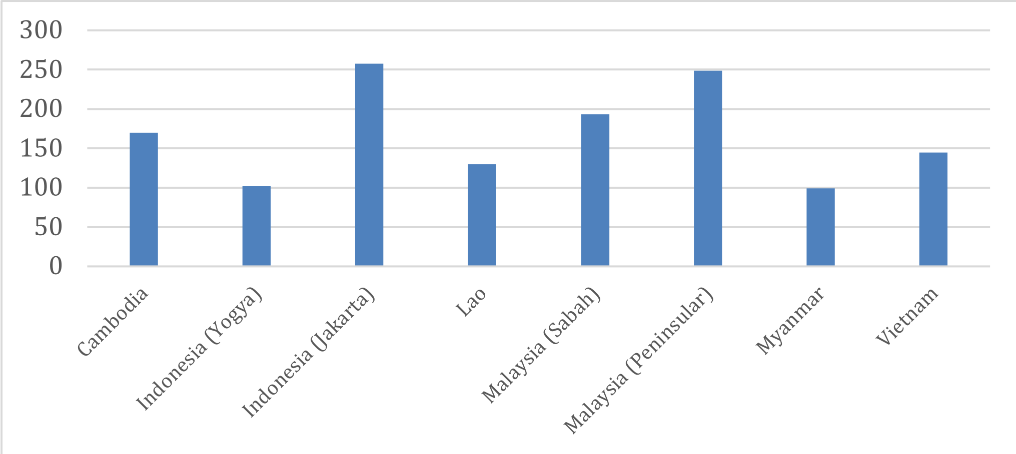


Figure 2.15: Minimum Wages Comparison in Selected ASEAN Countries, 2018 (in USD)
Source: ASEAN Secretariat, 2018

Gender Disparities

Gender disparities in labour force participation continue to persist. Female participation rate in the labour force is still relatively low; 50 per cent in 2000, increasing slightly to 53.4 per cent in February 2014 (ILO, 2014). According to the AIPEG (2017), the main drivers of low female participation are marriage, having children aged below 2 years in the household, low education attainment and changing economic structure, particularly the decline in the female-friendly sector of agriculture due to the transition from rural to urban areas. Meanwhile, Schaner and Das (2016) asserts that many Indonesian women exit wage work due to family and childcare constraints. JICA (2011) adds that many female Indonesians engaged in the informal sectors due to fewer job vacancies available for women in the formal sector and the flexibilities of working style in the informal sector.

The Indonesian government issued the Presidential Instruction No.9/2000 on Gender Mainstreaming in National Development, which regulates that gender mainstreaming is implemented in order to improve the role of women and to achieve gender equality in the family, the community, the state and the nation. To encourage the increase of women participation in employment, particularly in the formal sectors, the government produces regulations as a support system for female workers, such as the Labour Act No.13/2003, stipulating menstrual leave (two days per month), maternal leave (1.5 months pre- and post-childbirth, 3 months maternal leave in total), miscarriage leave (1.5 months), and provision of time for breastfeeding. In 2008, with the collaboration between the State Ministry for Women's Empowerment and Child Protection, Ministry of Manpower and Transmigration and Ministry of Health, a joint decree was enforced regarding the establishment of breastfeeding facilities in all buildings containing workplaces (JICA, 2011).

However, several studies find that females tend to have less working experience and working hours. For instance, when estimating the average working week based on industry and gender in Indonesia for 2000 and 2008 period, Klaveren *et al.*, (2010) conclude that the average hours females made are fewer than males, and that was the case in all industries, as presented in Table 2.3. Meanwhile, Taniguchi and Tuwo (2014) assert that working experience may also influence the gender wage gap. Women tend to have relatively shorter working experience as they enter and exit the labour market due to family considerations. Also, women anticipate shorter or more discontinuous work lives. In comparison to other countries, the gender wage gap in Indonesia in 2010 is 20.4 per cent. The gap is slightly lower than the gap in the US in 2012 (23 per cent), as shown in Table 2.5.

Table 2.3: Average Working Weeks by Industry and by Gender in Indonesia, 2000 and 2008

	total		men		women	
	2000	**2008**	**2000**	**2008**	**2000**	**2008**
agriculture	30.0	32.9	33.4	35.6	25.0	28.7
fishing	43.9	44.0	45.4	45.4	30.2	29.2
mining	43.8	44.4	45.7	45.6	35.0	35.2
manufacturing	43.1	43.8	45.9	46.5	39.2	40.3
utilities	43.1	44.0	43.3	44.2	41.0	42.0
construction	46.5	47.0	46.7	47.0	40.9	44.6
transport, storage, communication	50.9	49.1	51.0	49.9	48.3	42.6
wholesale and retail	46.5	49.4	48.4	49.6	44.0	49.2
hotels and restaurants	46.2	50.0	49.7	52.6	44.0	47.8
finance	42.5	43.7	42.9	45.0	41.5	41.0
real estate, renting, business	42.0	42.3	43.0	43.9	39.1	37.7
public administration, defence	41.1	41.7	41.6	42.4	39.3	38.5
education	34.3	34.5	36.1	37.1	32.2	32.5
health, social work	41.4	40.8	40.9	42.0	41.8	40.1
other community, social and personal services	41.0	43.1	44.7	45.4	35.1	37.3
private households	49.0	51.6	45.3	47.0	51.4	53.0
Total	**38.4**	**41.0**	**41.1**	**42.8**	**34.0**	**38.2**

Source: Klaveren *et al.*, (2010).
Notes: Average working weeks is in terms of hours per week.

Furthermore, Taniguchi and Tuwo (2014) find that the distribution of hours worked from those in the manufacturing sector tend to be longer than in the public sector. Meanwhile, the wage gap tends to be greater among public sector workers than those in the private sector. In terms of sector, the widest gender wage gap occurs in finance and real estate, followed by public administration. In fact, the public sector categories (*i.e.* public administration, education, and health and public services) tend to have wider gender wage gap compared to other industry categories, as shown in Table 2.3.

Table 2.4: The Gender Wage Differentials Based on Industry

Industry Category	All	Urban	Rural
Agriculture	28.1	26.5	28.3
Mining	26.1	28.0	25.5
Manufacturing	28.4	29.9	26.0
Construction and utilities	30.6	32.3	28.2
Retail and hotel	29.5	30.1	27.9
Transport and communication	31.8	32.4	29.5
Finance and Real estate	33.1	33.4	31.1
Public administration	33.0	32.9	32.5[a]
Education	32.9	32.7	33.3
Health and public services	32.3	32.3	32.6
Personal and household services	22.6	22.5	22.8
International corporation[b]	N/A	N/A	N/A

N/A = not available.
Notes:
[a] The linear regression of equation 2 is used instead of Heckman's two-step model because the model did not converge for rural areas.
[b] Results are insignificant.
28.1 means that the female worker's monthly real wage is 28.1% lower than the male worker's wage on average in the category.

Source: Taniguchi and Tuwo (2014).

Table 2.5: Selected Studies on the Gender Wage Gap

Country (Year)	Gender Wage Gap (%)	Study
Republic of Korea (2011) New Zealand (2011)	30.3* 4.2**	OECD (2014)
United States (2012)	23.0	AAUW (2014)
Indonesia (2010)	20.4	Fitrania (2013)
43-country average (various years) Indonesia (2008) Zambia (2005) Slovenia (2008)	18.4 13.7 46.0* 4.0**	Tijdens and van Klaveren (2012)
Philippines (2005) Management Administrative United States (1998)	 10.0 15.0 20.3	 Cabegin (2012) Blau and Kahn (2001)

AAUW = American Association of University Women, OECD = Organisation for Economic Co-operation and Development,
* (**) denotes having the widest (narrowest) gap across sample countries.

Source: Taniguchi and Tuwo (2014)

Ren and Miller (2012) argue that the gender difference occurs possibly because of the differences in the demand for, and supply of, education between males and females, greater positive self-selection of females into the labour force relative to males, a more limited supply of skilled female workers, different technological requirements between the female-dominated and male-dominated jobs, and discrimination against female workers that is less intense among the better educated.

2.2 Data

2.2.1 SAKERNAS and IFLS

There are two datasets related to the Indonesian labour market: (1) SAKERNAS (National Labour Force Survey) and (2) the Indonesian Family Life Survey (IFLS). SAKERNAS is provided by the government; the survey is conducted twice a year (in February and August). The data only cover labour market information and they are not longitudinal. The data can be obtained from Statistics Indonesia office or from their website. It is worth noting that the data are not freely accessible.

SAKERNAS has a bigger sample size and the size varies over time (compared to the IFLS). However, the consistency is still questionable. Many studies criticise SAKERNAS' quality and consistency, such as Manning (2006) who reviews the quality of SAKERNAS and finds that it is a reliable indicator of labour-force structure and long-term trends, but not of short-term trends. In addition, the questionnaire and variable definitions change occasionally. Thus, SAKERNAS has large year-to-year swings in their labour market indicators, which cannot be explained by any real events.

Meanwhile, IFLS data are collected by RAND Corporation in collaboration with some Indonesian institutions. RAND Corporation initiated the survey because there were no comprehensive and longitudinal data for Indonesia at that time (during 1990s). In general, the IFLS complements and extends the existing survey data available for Indonesia in some ways: as a longitudinal survey, the data are available for the same individuals from multiple points in time; thus, the IFLS data allows the opportunity to analyse the dynamic of any behaviours at the individual, household and family and community levels. Moreover, extensive research can be carried out regarding the living conditions of Indonesian households during this very tumultuous period, because the waves are conducted in a relatively long period; before, during, and after the Asian economic crisis in 1997/1998. In addition, the database is freely accessible. Specifically, for this study, IFLS data offer rich information on education and both public and private sector employments, which are the main variables in this study. Moreover, accessing IFLS data are simple; only requiring a registration in RAND Corporation's website, after which all IFLS data can then be downloaded freely. Considering those advantages, the IFLS data are selected for this study.

2.2.2 IFLS Data

The IFLS has had 5 waves so far; IFLS1 was fielded in 1993, IFLS2 in 1997, IFLS3 in 2000, IFLS4 in 2007-2008, and IFLS5 in 2014-2015. There is also IFLS2+ which was carried out in 1998 with 25 per cent of the sample; mainly aiming to measure the immediate impact of the Asian economic crisis in 1998 (the details of the IFLS information are shown in Table 2.6). Because the IFLS are a longitudinal survey, the earlier IFLS wave drew their sample from the previous waves. For instance, IFLS5 drew their sample from IFLS1, IFLS2, IFLS2+, IFLS3, and IFLS4.

Table 2.6: The IFLS Survey Summary

Wave	Fielding Period	Survey Size		Implementers	Contact/Re-contact Rates
		Households	Individuals		
1	1993	7,224	22,000	RAND and Lembaga Demografi (University of Indonesia)	93 per cent of selected households contacted
2	1997	7,619	33,934	RAND, UCLA, and Lembaga Demografi (University of Indonesia)	94.4 per cent of IFLS1 households re-contacted
2+	1998	2,055		RAND, UCLA, and Lembaga Demografi (University of Indonesia)	94 per cent of individuals not in IFLS2; 96 per cent of IFLS2 respondents
3	2000	10,435	43,649	RAND, and Population Research Center (University of Gadjah Mada)	95.3 per cent of IFLS1 households re-contacted
4	2007/08	13,535	44,103	RAND, Centre for Population and Policy Studies (University of Gadjah Mada), and Survey METRE	93.6 per cent of IFLS1 households re-contacted.
5	2014/15	16,931	58,337	RAND, and Survey Meter	92 per cent of IFLS1 households re-contacted.

Source: The IFLS User Guide, various editions.

IFLS1 was conducted based on the sample representing about 83 per cent of the Indonesian population living in 13 provinces (out of 27 provinces) in 1993. Those provinces were selected to maximise the representation of the population, capture the cultural and socioeconomic diversity of Indonesia, and be cost-effective to the survey given the size and terrain of the country. Within each of the 13 provinces, enumeration areas (EAs) were randomly chosen from a nationally representative sample frame used in the 1993 SUSENAS, a socioeconomic survey of about 60,000 households. The IFLS randomly selected 321 enumeration areas in the 13 provinces, over-sampling urban EAs and EAs in smaller provinces to facilitate urban-rural and Javanese–non-Javanese comparisons (Strauss, *et al*., 2016).

Furthermore, there are two different questionnaires for the individual-household and community levels. This study will focus on the data of individual-household levels, particularly the adult respondents, since all the data required in the present study are available in the survey for primary respondent. The re-contact and completion rates of IFLS1-IFLS5 are provided in Appendix I.

In terms of data collection, the IFLS data are collected by the means of face-to-face interviews using paper and pencil questionnaires. From 1997 onwards, each interviewing team was accompanied by a team of data editors; the team entered the data on laptops and performed extensive consistency checks on the same day the interview took place. If there were any inconsistencies, it was corrected in the field, and by revisiting respondents when necessary. Afterwards, electronic data were sent to the Principal Investigators to facilitate monitoring of the data quality during the fieldwork.

The respondents' selection follows the rules: if an entire household, or target respondent(s) moved then they were tracked as long as they still resided in any one of the 13 IFLS provinces, irrespective of whether they moved across those provinces. Target respondents were individuals who split off into new households, provided they were a main respondent in 1993 (which means that they were administered one or more individual questionnaires), or they were born before 1968 (*i.e.* they were 26 years old or older in 1993). Not all individuals were tracked in order to control costs.

There is a significant increase in the number of households as IFLS3 expands the rules for following households who moved out, *i.e.* (1) 1993 main respondents; (2) 1993 household members born before 1968; (3) individuals born since 1993 in the original 1993 households; (4) individuals born after 1988 if they were resident in an original household in 1993; (5) 1993 household members who were born between 1968 and 1988 if they were interviewed in 1997; and (6) 20 per cent random sample of 1993 household members who were born between 1968 and 1988 if they were not interviewed in 1997.

Furthermore, the rules for interviewing individuals were also expanded. In the original IFLS1 households, every individual could be interviewed or be given a proxy interview, whether or not they had been a household member in IFLS1. In split-off households, all IFLS1 household members, their spouses and biological children, were to be interviewed, but not others (not just the target respondents for tracking, but also their spouses and children, as in IFLS2).

The re-contract rates are relatively high, more than 90 per cent, for example IFLS5 the dynasty recontact rate was 92 per cent. For the individual target households (including split-off households as separate) the re-contact rate was a little lower, 90.5 per cent. According to Strauss *et al.* (2016), these re-contact rates are as high as or higher than most longitudinal surveys in the United States and Europe. High re-interview rates were obtained because Rand was committed to tracking and interviewing individuals who had moved or split off from the origin IFLS1 households. High re-interview rates contribute significantly to data quality in a longitudinal survey because they lessen the risk of bias due to non-random attrition in studies using the data.

The present study will elaborate the IFLS data in each chapter (Chapter 3 – Chapter 5), based on specific questions on labour and education. Each chapter (Chapter 3-5) also provides sample restrictions, summary statistics and additional information on robustness or sensitivity tests.

Chapter 3 The Return to Education in the Waged Sector in Indonesia, 2000 and 2014 Periods

3.1 Introduction

Education in Indonesia has continuously experienced rapid expansion, in particular university and senior high school, which is reflected by the Gross Enrolment Rate (GER) data, as shown in Figure 2.2. This is possibly because higher education is the most important factor in career progression, especially in the formal sector; and several important factors that influence this expansion are the government policies such as the nine-year compulsory education, as explained in Chapter 2.

Theoretically, the human capital theory suggests that wage increases in line with education level. Although wages do not reflect market wages in public sector[9], estimating private return to education is useful regarding the incentives set by the government to invest in education. Thus, the recent development of return to education is always needed to provide the latest information both for individuals and for the government.

Most empirical studies for Indonesia's case confirm that the (private) return to education increases in line with the level of education (such as: Comola and de Mello, 2010; Gropello and Sakellariou, 2010; Dumauli, 2015; and Dong, 2016). Regarding the change in recent years, a supply and demand framework suggests that the increase in participation rate of education increases the labour supply; thus, there is an outward shift in labour supply in the long run. In the short run, the opposite effect may occur; it could withdraw potential workers for studentship from the labour force. On the demand side, there is also an increase number for educated workers due to technical and economic progress.

For the change over time in the return to education, the empirical studies' result varies according to the method, period, and context of study. Comola and de Mello (2010) find most of returns to education level in Indonesia increase in general, between 1996 and 2004. In contrast, Gropello, E.D. Sakellariou (2010) find a decrease in return to education in Indonesia, particularly for general, vocational high schools and universities (relative to primary schools) between 1994 and 2007. Slightly different, Purnastuti *et al*. (2013) find

[9] Before the bureaucratic reform, wage mechanism in public sector in Indonesia is based on cronyism and political patronage. After the reform, the queues of potential employees occur, as explained in Section 2.1.

the returns to education in 2007/08 generally lower than the returns in 1993, except for the return to university, which increases between those periods, for both males and females.

This chapter elaborates on the effect of education expansion on wages, particularly the financial return to education in the waged sector (public and private sectors), following the full implementation of several important education policies. Thus, this study employs the years of 2000 (IFLS3) and 2014 (IFLS5). 2000 represents the situation before the Education Reform based on the Law No. 20 of 2003 and Higher Education Law No. 12 of 2012 was implemented, and 2014 represents the situation after those policies were fully implemented. Meanwhile, education level (*i.e.* primary school or below, junior high school, senior high school and university) is chosen to simplify the analysis and effectively distinguish the return to education between different education levels (the particular interest is in higher education level). Moreover, this allows non-linear effects of the level of education (Steiner and Wagner, 1996); this is also in line with the policy implementation that is commonly based on education level. Furthermore, waged sector (only public and private sectors) is chosen considering the data availability.

The aims of this study are: to investigate how return to education varies across education levels and across time (between 2000 and 2014 period), to provide a more systematic model specification of return to education in Indonesia (including religion and ethnicity variables that rarely appear in the previous research), and to provide an extended analysis by separating the sample by gender and sector.

This study is motivated to estimate and update the return to education in different levels of education and to investigate whether the large increases in education participation which have occurred in the last decades affect the return to education in Indonesia.

Therefore, the research questions of this study are: (3.1) What are the estimated returns to education in 2000 and 2014? (3.2) Do gender and sector affect the return to education? and (3.3) How does the return to education change between those periods? To answer these research questions, IFLS3 (2000) and IFLS5 (2014) will be used, and the Ordinary Least Squares (OLS) method will be used to estimate the return to education and the change, and the robustness tests related to endogeneity and selection issue will be carried out to further corroborate the results.

The structure of this chapter is as follows: Section 3.2 briefly reviews the theoretical framework, which consists of:

- the human capital theory,
- the signalling theory,
- Mincer wage equation and its main methodological issues,
- the supply and demand framework of labour market (to analyse the change in education participation); and,
- previous empirical studies.

Furthermore, Section 3.3 discusses the method and data. Section 3.4 presents the estimation results from IFLS for 2000 and 2014 periods. And finally, Section 3.5 gives conclusion of the return to education in waged sector.

3.2 Literature Review

This part addresses some important questions: firstly, why is there a return to education? It will be answered by considering the human capital and signalling theory. Secondly, how is it usually measured? The standard Mincer wage equation will be used to answer the question. Thirdly, what is the estimated rate of return to education? It will be explained by some empirical studies in the US, the UK, Indonesia and other similar countries, and based on regions as well. Then, what is the likely impact of the expansion? The simple supply and demand framework of higher education will address the question. And finally, what has happened to the rate of return to education? This question also, similarly, will be explained by some empirical studies in the US, the UK, Indonesia and other similar countries.

3.2.1 Why is there a Return to Higher Education?

The accumulation of human capital is perceived as an investment decision in the standard economic model (Schultz, 1961; Mincer, 1974; and Becker, 1993). During a period of education, an individual gives up some proportions of wage and spends on education costs (both direct and opportunity costs) in exchange for higher future wage. Therefore, education must yield a higher rate of return compared to other investments such as working, in order to be pursued from an economic perspective.

Theoretically, the more education individuals acquire, the better they are able to take in new information, skills and technologies; consequently, their productivity increases (Poteliene and Tamasauskiene 2013). To explain the return to education, there are several alternative theories that are commonly used, *i.e.* the human capital theory and the signalling theory. The main difference between those theories is whether education endows or acts just as a signal; indeed, this difference could affect education policies. However, both theories agree that there is a positive relationship between education attainment and return to education (wages). The detailed explanations of both theories are as follows.

Human Capital Theory

The Human capital theory, pioneered by Gary Becker, is fundamentally based on the belief that the role of employees in production process is similar to the role of machinery and other forces of production. As a result, the aim of investing in higher education level is to provide essential knowledge, skills, and abilities for employees and is evaluated in the same way as investment in facilities and equipment for the purpose of improving productivity[10].

Becker developed the first model to measure the returns to education, *i.e.* the internal rate returns or IRR (symbolised as R_X); this concept is based on workers' point of view. The internal rate of return is the rate of discount that equates the net income stream with and without an investment. For example, an individual considers higher education as an investment, the internal rate of return is the present value (PV) of additional income compared to those who had the right to, but did not pursue higher education to the present value of cost (opportunity cost through foregone earnings and, under a private financing scheme, the direct cost of study). If this rate of return is higher than the market interest

[10] Acemoglu (2014) divided sources of human capital into 5 categories: (1) innate ability, (2) schooling or education, (3) school quality and non-schooling investments, (4) training, and (5) pre-labour market influences. And all of those categories assume that employees have different amounts of skills or human capital.

rate at which the individual can borrow loans, education represents a worthwhile investment for the individual (Potelienė and Tamašauskienė, 2014).

The model assumes that certain education level is chosen to maximise the expected present value of the stream of future wage, until retirement at date T, and education in this case is measured by years of schooling. Moreover, individuals spend costs during higher education period, in terms of net cost of education (C_X). At the optimum level of schooling, the PV of the returns to higher education is equal to the cost of higher education. Thus, the equilibrium condition is characterised by:

$$\sum_{t=1}^{T-X} \frac{W_X - W_{X-1}}{(1+R_X)^t} = W_{x-1} + C_X \quad (3.1),$$

where:

R_X: the internal net return to higher education; X is years of schooling (level of education), X-1 is lag of year (level) of schooling; W_X: wage with X years of schooling.

The optimal level of investment in higher education occurs when the internal net return is higher than the market rate of interest (r), or $R_X > r$, assuming no borrowing constraints. If T is large enough, the left-hand side of the equilibrium condition can be approximated, thus the equilibrium condition is:

$$\frac{W_X - W_{X-1}}{R_X} = W_{X-1} + C_X \quad (3.2),$$

If C_X is sufficiently small, then the equilibrium can be arranged as:

$$R_X \approx \frac{W_X - W_{X-1}}{W_{X-1}} \quad (3.3),$$

$$R_X \approx log\ W_X - log\ W_{X-1} \quad (3.4),$$

where ≈ means 'approximately equal to'. The equation above implies that the return to higher education is approximately the difference in log wage (W) between X_t years of schooling and its lag variable (X_{t-1}), or the variation of log wage (W) with years of schooling.

Becker (1994) also carries out some empirical studies to apply the theory which estimates the rates of return of white male with higher education qualifications in the US for 1939 and 1949, using the national data on the income of persons at different education levels from the 1940 and 1950 censuses. The study calculates the monetary returns from attending higher education by comparing the returns and costs in absolute value. In addition, the rate of return to a cohort can be calculated from the stream of total (cohort) absolute differentials or from the mean (per capita data) differentials. The returns and costs of the cohort are adjusted by mortality, growth and taxation. The study finds that the best single estimate for the private rate of return for 1939 cohort is 14.5 per cent. On the other hand, the private rate of return for 1949 cohort is 12.7 per cent if income differential grows at 1 per cent per annum, or about 1 per cent lower or higher if the income differential grows at 2 per cent; and the best single estimate is around 13 per cent.

However, there are some criticisms aimed towards Becker's theory, particularly some empirical challenges. Firstly, the analysis puts aside the difference of education quality as one of the important issues. Secondly, the data limitations may lead to overestimation of the returns because the analysis would be done with cross-section sample, and it does not follow the lifetime returns of the same age cohort. And finally, the return to human capital is hard to specify and measure, as one would find oneself dealing with average benefit instead of specific returns to each individual investment (Teixeira, 2014). Heckman *et al.* (2016) also note some challenges of the IRR: (1) it requires lifetime earning profiles; (2) earnings are observed only at schooling level selected by agents, thus observed earning profiles are subject to selection bias; (3) it is difficult to calculate non-market benefit and non-pecuniary costs.

Despite such limitations in measurement; the human capital theory is still important and will become the basis theory for further analysis. Becker (1964) argues that there are some important assumptions in the human capital theory: productivity is an increasing function of the human capital level of the worker. Human capital includes not only formal education, but also experience and on-the-job training. In human capital model (as under perfect competition), labour is paid the value of its MP and assumed wages are determined by the workers' educational attainment, experience and training. Subsequently, the human capital model implies that worker's characteristics, or the supply side, determine earnings and it is only through exogenous shocks that the demand side impacts real wages.

In the development of human capital theories, Blundell *et al.* (2000) classify the return to investments in education into three types, *i.e.* (1) private financial return to education, mainly that acquiring education improves the wages and/or employment prospects of an individual; (2) private non-financial returns to education, comprising improvements in an individual's welfare, which is not only wages but also better working environment and other non-financial measurements; and (3) social return to education, including benefits to other individuals and society[11]. The present study focuses on the first category, *i.e.* wage premium.

Finally, Mincer Wage Equation, a common single-equation model, is used to measure wage in relationship with education level. The equation, named after the proposer, Jacob Mincer, is very popular because it is based on a formal model of investment in human capital (Lemieux, 2006). The principle of Mincer equation is derived from the Human Capital theory; wages are determined by the workers' educational attainment and experience (excluding training).

Signalling Theory

The signalling theory offers a sensible economic explanation for the observed positive correlation between wages and education levels. Nobel laureate Michael Spence originally identified the theory in 1973. Afterwards, Kenneth Arrow and Joseph Stiglitz developed the model. According to the theory, highly educated employees are likely to be more productive, not only because they are better trained, but also because they possess some natural skills required to get the educational degree. If these individuals' characteristics are to persist when working, educated employees should also be easier to train and firms will obtain higher revenue from them. Therefore, employers could distinguish high-productive workers just by observing the signal given by the attainment of different education levels (Livanos and Núñez, 2012).

The main assumption of the signalling model is the asymmetrical information between the employers and the employees regarding the employees' true skill level, assuming that the employers were risk neutral. In this case, employers expect to hire capable employees for better wage positions; and the signalling mechanism can play an important role in accurate job-market matching. The model also assumes that signalling costs are negatively correlated with the employee's productive capability, implying that high

[11] Such as: lower crime, greater social cohesion, trust and tolerance, political stability, greater social mobility, greater social capital, better educational parenting, more likeliness to vote and support democracy better life satisfaction better health quality, and longer life expectancy. It could also lead to increasing tax revenues, faster economic growth, greater innovation and labour market flexibility, increased productivity of co-workers, and reduced burden on public finances from better coordination with other social policy areas such as health and crime prevention.

ability individuals will have lower cost of signalling, while lower ability individuals will have higher costs.

An individual is assumed to select signals, for instance high productivity employees tend to choose to obtain higher education qualifications, and low productivity employees tend to choose to obtain lower level of education such as senior high school or lower; thus maximising the difference between offered wages and signalling costs. An individual will also invest in higher education if there is sufficient return as defined by the wage schedule (including other returns to education and signals from consumption goods).

There is also an informational feedback in the job market to the employers over time, as shown in Figure 3.1. As new market information comes into the employers throughout hiring and subsequent observation of productive capabilities related to signals, the employers' conditional probabilistic beliefs are adjusted, and the new round starts. Meanwhile, the wage schedule facing the new entrants in the market commonly differs from that facing the previous group of employees. The system will be stationary if the employers start out with conditional probabilistic beliefs that after one round are not confirmed by the incoming data that are generated, such beliefs are referred to as self-confirming (Spence, 1973).

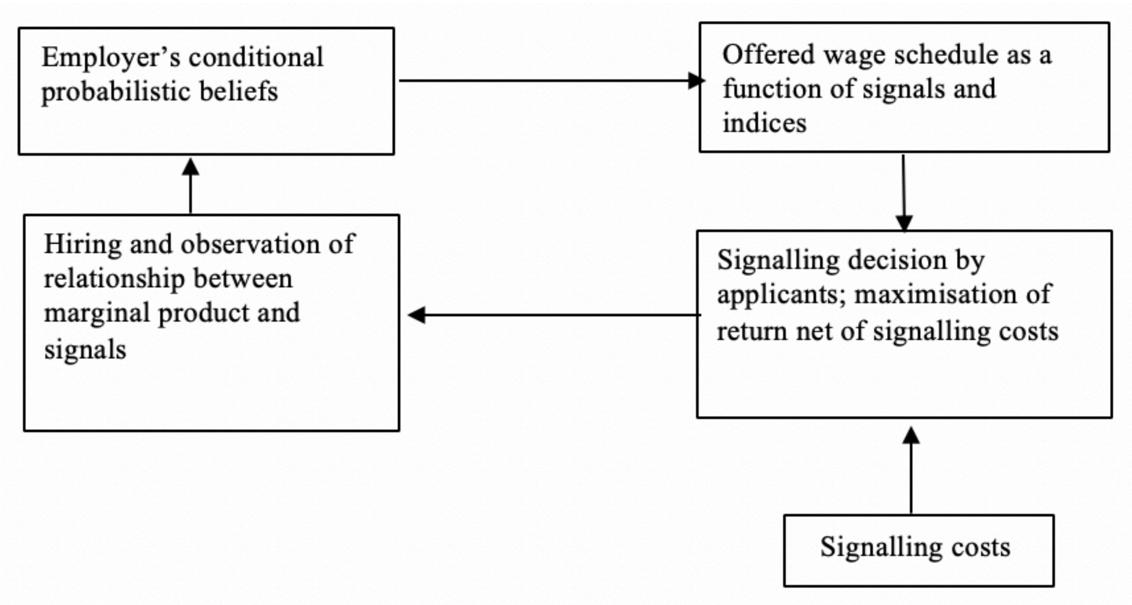


Figure 3.1: Informational Feedback Loop in the Job Market
Source: Spence (1973).

Even though it is a loop, the equilibrium could exist (the equilibrium is defined as a set of components in the cycle that regenerate themselves). As a result, the employer's belief would be self-confirming, or offered wage schedules would regenerate itself, or applicant's behaviour would reproduce itself in the next round of the loop. The occurring equilibrium is separating equilibrium (E^s), which requires the wage schedule to induce self-selection, as explained above. To illustrate this equilibrium, Spence divides the population into two groups facing one employer, as summarised in Table 3.1. The signal under consideration is education[12], measured by an index y and is subject to individual choice.

Table 3.1: Spence's Signalling Model

Group	Marginal Product	Proportion of population	Cost of education level (C_X)
I	1	Ω_1	X
II	2	$1-\Omega_1$	X/2

Source: Spence (1973).

Where: Ω_1 is the proportion of group I; if the employer believes that there is a level of education X* below which productivity is 1 and above which productivity is 2. The wage schedule w(X) offered will be:

$$w(X)\begin{cases}1, & X < X* \\ 2, & X > X*\end{cases} \qquad (3.5),$$

members of each group will select optimal levels of education, or as depicted by the green line in Figure 3.2. Consider the person who will set X<X*, then their optimal level of education is X=0 because education is costly; until he/she reaches X*, there are no benefits to increase X, given the employer's hypothesised beliefs. For a person who sets X≥X*, their optimal level is X=X*, because further increase would merely incur costs with no corresponding benefits. Thus, workers will set either X=0 or X=X* in equilibrium, as depicted in Figure 3.4. Given the employer's initial beliefs and the fact just deducted, if employer's beliefs are to be confirmed, then workers in group I must set X=0, and workers in group II must set X=X*.

In the case where the employer's beliefs are confirmed, a signalling equilibrium occurs. The model defines the conditions on behaviour by two groups, in order for the employer's beliefs to be confirmed, in algebraic, as follows: (1) There is no rational reason for

[12] Education costs are both monetary and psychic.

someone to choose a different level of education from 0 or X*; (2) group I sets X=0 if 1>2-X*, which implies that the return for not investing in education is higher than investing in education; (3) group II sets X=X* if 2-(X*/2)>1, which implies the return for investing in education is higher than not investing in education; (4) putting conditions (2) and (3) together, if 1<X*<2, then employer's initial beliefs are confirmed; (5) in interval [1,2], there are infinite equilibrium values of X*, but those values are not equivalent from a welfare point of view. The higher X* the worse off is Group II, while Group I is unaffected; and (6) if there is no signalling, workers will be paid of this unconditional expected marginal product $\Omega_1+2(1-\Omega_1) = 2-\Omega_1$; thus, group I is worse off than it was with no signalling at all. If the proportion of group I is 0.5, group II may also be worse off than it was with no signalling; in such a situation, everyone would prefer no signalling.

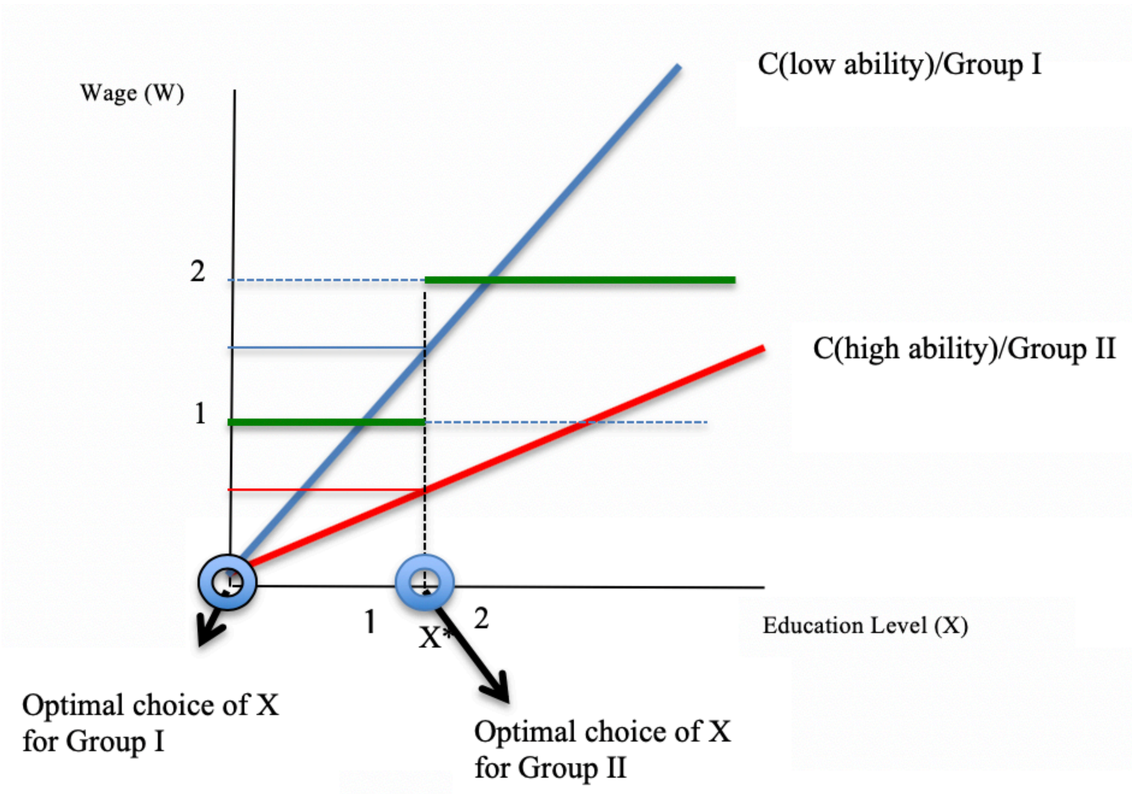


Figure 3.2: Simple Signalling Framework
Source: Spence (1973).

The signalling model has an implication that even if education has no real contributions to the marginal product of the worker, the combination of the beliefs of the employer and the presence of signalling transform the education level X* in a prerequisite for the higher paying job. In other words, education could increase the marginal product of labour, without this necessarily being true.

Regarding the empirical research, Tyler *et al.*'s finding (2000) supports the signalling theory. They estimate the signalling value using the General Educational Development (GED) equivalency credential and Social Security Administration (SSA) earnings data. The study compares the wages of individuals who had the same GED test scores in the US with different passing standards and assumed that individuals who had the same GED score acquires equal amount of human capital and have the same productivity level.

The study finds that the GED signal increases the earnings of young white dropouts by 10 to 19 per cent and there is no statistically significant effect for minority dropouts. Although there are many evidences of higher education as signalling, the validity test is needed to criticise; as Page (2010) concludes that empirical tests of signalling model were most prevalent during the late 1970s and 1980s, but many were of questionable validity. Since the late 1980s, there have been fewer studies that have attempted to identify the role that signalling plays in the labour market, but the literature has laid more emphasis on finding external sources of identification.

Furthermore, many studies have tried to prove whether education is just as a signalling or more than that (human capital). Most empirical studies find that higher education supports the human capital theory more, as Harmon *et al.* (2000) find that there are some significant effects of ability on wages, and education plays a largely productivity-enhancing role. The study uses two databases, *i.e.* the GB National Child Development Survey[13] (NCDS) and the International Adult Literacy Survey (IALS) which record wages and ability data for the UK. They apply three approaches to finesse the problem of education, either as signalling or as human capital. Firstly, they attempt to control ability and examine if education still has a strong influence on wages, any difference could be attributed to the signalling value of education. Secondly, they compare estimated returns that controls for ability with those that do not. And finally, they distinguish between ability and productivity to directly include ability measures. However, there is a problem with this method: the ability measures need to be uncontaminated by the effects of education; otherwise, they will pick the productivity-enhancing effects of education.

Similarly, Chevalier *et al.* (2002) conclude that education in the UK strongly supports the human capital explanation. They study the effect of education on wages, whether it is

[13] NCDS is a cohort study of all individuals born in the GB in a particular week in 1958 whose early development was followed closely and whose subsequent careers have been recorded including their wages.

because education increases productivity or because of a signal of ability, using the large Labour Force Survey (LFS) data pooled from 1993 to 2001 from England and Wales. The approach used to distinguish between the two theories allows for the possibility of employer learning. The study finds that the effect of education on wages is relatively large – possibly approaching 10 per cent per additional years of education.

From the explanation of both the theories and empirical evidence of human capital and signalling above, the key difference between signalling and human capital models is that signalling models allow firms to draw inferences about unobserved characteristics of employees. Those inferences can be based on schooling or work experience of employees, or on direct measures of some aspects of job performance (Weiss, 1995). On the other hand, the human capital theory argues intuitively that education endows an individual with productivity-enhancing ability, and this increased productivity results in increased wages in the labour market. Human capital theory follows the competitive market theory, which requires labour to receive a wage equal to their marginal product. Meanwhile the signalling theory offers an oppositional argument, which holds that education only reflects inherent abilities. These inherent abilities (not education itself) increase productivity and lead to higher wages (Kjelland, 2008).

Although there is a significant difference between the human capital and the signalling theory, and this could have different policies' implications, both theories agree that wage premium is an increasing function of education attainment.

3.2.2 How is the Return to Education Usually Measured?

There are some frameworks of the Mincer wage equation: individuals have identical abilities and opportunities, credit markets are perfect, the environment is perfectly certain, but occupations differ in the amount of schooling required. Individuals sacrifice wages while in school but incur no direct costs. Because individuals are *ex ante* identical, they require a compensating wage differential to work in occupations which require a longer schooling period. And finally, the compensating differential is determined by equating the present value of earnings streams net of costs associated with different levels of investment (Heckman *et al.*, 2006). Furthermore, the Mincer wage equation assumes that (1) an individual with X years of schooling has wages which do not depend on age, for example a 40 year old new graduate will obtain the same as an 18 year old new graduate, when they are graduated in the same time; (2) present value of lifetime income is the same across individual regardless of schooling if no post-schooling investments are made; and (3) the number of years spent at work are independent of the number of years schooling (Bunzel, 2008).

In the regression, the standard Mincer wage equation is represented by log of wage (W), which only depends on education (years of schooling/X), years of experience (L), and the squared of experience (L^2), with a linear relationship between education and wages, the proof of log linear relationship is shown in Appendix II. In the present research, Mincer wage equation for panel data is used to analyse the change of wage for a given sample individuals (i) over time, the standard Mincer wage equation is as shown below:

$$\ln W_{i,t} = \beta_{0,t} + \sum \beta_{1,n,t} X_{i,n,t} + \beta_{2,t} L_{i,t} + \beta_{3,t} L_{i,t}^2 + \varepsilon_{i,t} \qquad (3.6),$$

where:

$ln\ W_{i,t}$: wage of individual i at time t (in log),

$X_{i,n,t}$: education level of individual i at time t, n is 1 to 4 (primary to university level) or n =1 if education variable is defined as years of schooling,

$L_{i,t}$: years of experience of individual n at time t; or potential labour market experience,

$\varepsilon_{i,t}$: a random error term,

$\beta_{1,n,t}$: the rate of return to education for n (1…4 level of education) at time t, β_2 and β_3 are parameters of experience, $\beta_{0,t}$: the intercept, Σ is set (vector) of explanatory variables; n is individual (i=1…I) and t refers to time period.

In an empirical research, actual or potential wage (W) can be used in the estimates. Actual wage refers to wages that individuals actually receive. If actual data are not available, an alternative variable that can be used is potential wage, *i.e.* the potential income that might be earned if the individual *played by the rules* and worked for salary as much as other individuals do. There are some wage proxies: net or gross, hourly, weekly, monthly or yearly wages (Pereira and Martins, 2001). Swaffield (2000) defines wage as the gross average hourly wage and is transformed into the real wage by using a nominal wage index (constructed from the New Earnings Survey's average hourly wage). Meanwhile, Chevalier *et al.* (2002) calculate an hourly wage rate from the ratio of usual wage including overtime pay to usual hours (from the respondent's main job including overtime) and deflate all wages to 1993 prices using the Retail Price Index. One advantage of using hourly wage is that it could eliminate unobserved heterogeneity caused by the omitted working hours (Li and Urmanbetova, 2007); moreover, workers sometimes have more overtime in one month, but not that much time in another month. Transforming yearly wage into hourly wage is hoped to eliminate this problem.

Education attainment (X) could be defined as the highest level of education successfully completed and is either indicated by the highest educational qualification (vocational or academic) achieved, or by the number of years of education or schooling completed or in which case each year is regarded as a kind of level (UNESCO, 2010). Most studies calculated this variable by attaching the average of years to several standardised education levels or the total years of schooling (Asplund and Pereira, 1999). Meanwhile, other studies used a set of dummy variables as individual's educational and vocational qualifications, allowing non-linear effects of the level of education. This specification also takes into account that, for a given completed educational/vocational degree, fewer rather than more years are considered as a positive signal (Steiner and Wagner, 1996).

In terms of labour market experience (L), the standard Mincer wage equation also indicates a linear function between experience and wages, due to homogeneous individuals. Meanwhile, the quadratic function of experience variable could capture the fact that on-the-job training investments decline over time, as stated in the standard lifecycle human capital model, the algebra of this relationship is provided in Appendix III. Most empirical studies used actual and potential experience as labour market experience variables, indeed potential experience can be used if the data were not available. There is a distinction between both variables; actual experience refers to the sum of lifetime hours spent working and training while potential experience refers to the time elapsed since leaving school (Regan and Oaxaca, 2009). Thus, there is a possibility of different data between actual and potential experience.

The equation postulates that experience in the labour market has a positive impact on wages. Furthermore, the effect of the square of experience is negative which implies that there are diminishing returns in experience. If the model assumes a linear relation between wage and experience, the correlation between both variables shows a concave shape. In addition, to reflect labour market experience, some studies also used job tenure which represents the individual's years of experience in their present job; this variable is usually viewed as a measure of firm-specific training and knowledge, at the same time. The hypotheses are the same as experience variables.

Burdett and Coles (2010) argue that wage changes both with experience and with tenure because of two reasons: (1) individuals accumulate human capital by working, for example: typists become better typists while working as typists and; (2) human capital can be dichotomised into general human capital and firm-specific human capital. A worker who enjoys an increase in general human capital becomes more productive at all jobs (related to experience), and accumulating firm-specific human capital implies a worker is only more productive at that firm (tenure effect). As such, workers who change jobs, or those who are laid off, lose their firm-specific human capital, but keep their general human capital. As experience variable, the quadratic function of experience variable could capture the fact that on-the-job training investments decline over time.

In the development, most research modifies the standard Mincer wage equation with control variables, *i.e*. personal characteristics such as sex (Comola and de Mello, 2009), and marital status (Chevalier, *et al.*, 2002; Comola and de Mello, 2009); job and firm related variables such as present labour market experience or tenure (Purnastuti *et al.*, 2013), firm size and firm age (Pereira and Martin, 2001), industries (Comola and de Mello, 2013), formal and informal sectors (Dasgupta *et al.*, 2015); urban and rural area residential (Dumauli, 2015); and some interaction terms such as gender-material status and gender-dependency ratio (Comola and de Mello, 2009). These modifications are done with the aims to capture other factors that may affect the wage equation.

It is worth noting as well that Mincer (1958) also asserts that the resulting *age-wage profile* was steeper for more educated employees than for those less educated. In other words, log wage is not a strictly separable function of education and age. There is no such thing as a single rate of return to education but rather a different rate of return for each age group. In contrast, Mincer (1958) also points out that in schooling, experience and wage, the experience-wage profiles are relatively parallel for different education groups.

Limitations of the Mincer Wage Equation

The Mincer wage equation has been commonly used to estimate return to education, and also offers a good starting point (Humphreys, 2012). Lemieux (2006) asserts that Mincer wage equation provides a parsimonious specification that fits the data remarkably well in most contexts. However, there are some valid critiques of the model. In terms of assumptions, mainly the same present value of lifetime incomes cannot be fulfilled, only for simplification purpose.

There are some limitations of the Mincer wage equation, *i.e*. (1) an endogeneity of education exists due to an ability bias and other omitted variable bias (this will be explained in detail in the next section); (2) measurement error in education variable causes the biased and inconsistent OLS estimates. Bias due to measurement error in the schooling variable is generally known to produce an attenuation bias in the coefficient of schooling; (3) sample selection bias occurs due to non-random selection of the sample used for the estimation process, when the sample is only based on a subpopulation; (4) the relationship between education and wages could be non-linear, this may be because of *sheepskin* effects, where achieving the final credential (*e.g.* a high school certificate or university degree) is more important than non-credentialed education. For instance, completing four out of four years of a higher education degree may well result in a large wage premium, but a person who completes only three out of four years of the same degree may receive a much smaller wage premium, as Mincer (1997) argues; (5) there are differences in experience profile in the labour market or heterogeneous experience, for example: high-skilled jobs may include a significant amount of “on the job” training and greater opportunities for professional advancement, while low-skilled jobs could only have little experience (Humphreys, 2012; Firpo *et al.*, 2005)). The present study will focus on two of its limitations, *i.e*. the endogeneity and sample selection bias, which will be discussed in the next part.

Although there are many limitations of Mincer wage equation, Lemiux (2003) concludes that the equation is still a good approximation in many cases, but it may overstate or understate the effect of experience and schooling on wages for some groups. Lemiux (2003) evaluates the empirical performance of the standard Mincer Wage Equation for the US data, using the Current Population Survey (CPS) for the years 1979-2001. The research finds that the equation does not appear to fit the data of the US in the 1980s and 1990s, but it fits the data from the 1960s and 1970s; because there was an increasingly convex function of years of schooling and experience-wage profiles are no longer parallel for different education groups. However, the Mincer wage equation remains useful and accurate in a stable environment where educational achievement grows smoothly over successive cohorts of employees. In short, the Mincer wage equation remains a parsimonious and relatively accurate model of the relationship among wage, education and experience, particularly in a stable environment.

Endogeneity Issue in Mincer Equation

Referring to the first assumption of the Mincer wage equation: all individuals are identical apart from the difference in education and training, it seems that the assumption cannot be fulfilled since different people cannot be identical, as they have different social environment, family background *etc*. Endogeneity problem arises due to unobservable variables such as ability. Ability can determine wage in the labour market and at the same time it can be correlated with education. This problem leads to biased results in the OLS estimations, because the education variable will be correlated with the error term in wage equation. The evidence of this endogeneity problem is explained in Appendix IV.

To address the endogeneity issue, instrumental variable method can be used to solve this problem. The first step is estimating the predicted value of schooling variable, *i.e*.

$$\hat{X} = \beta_0 + \sum \beta_{1,n,t} K_{i,n,t} + \beta_{2,n,t} \sum Z_{i,n,t} + \varepsilon_{i,t} \quad (3.7),$$

where: $\hat{X}$ is predicted value of education/schooling variable; $K_{i,n,t}$ is vector of all explanatory/control variables; and $Z_{i,n,t}$ is instrument variables.

After the predicted value of education is obtained, in the second stage, education variable is replaced by the predicted one (from the first stage):

$$\ln W_{i,t} = \beta_0 + \sum \beta_{1,n,t} K_{i,n,t} + \beta_{2,t} \hat{X}_{i,t} + \varepsilon_{i,t} \quad (3.8).$$

In terms of instrumental variables, some conventional variables can be used, such as:

1. family background, including the parent's education, and household wealth (Blackburn and Neumark, 1993). The idea is more educated families affect their wage by providing education friendly environment or by providing more financial help for their children. However, this could become an unfit instrument if family factors affect the wage directly by securing better and well-paid jobs for the children using their social affiliations and power;
2. IQ or other academic scores (Harmon *et al*., 2003), those variables are assumed to capture natural ability; students with greater abilities (or some other hidden advantages) are likely to receive more schooling and also receive higher incomes, which could result in a correlation between schooling and wage that does not describe a causal link. If ability is related to both schooling and wage, then the

standard Mincer equation would give an upward biased result, and will also cause a convex relationship between education and (log of) wages (Humphreys, 2012);

3. siblings or twin data, such as Butcher and Case (1994) use of "the presence of any sisters" within a family as an instrumental variable for schooling of female workers on the basis that gender composition of siblings in a family has a significant effect on educational attainment but no effect on inherent ability;
4. availability of educational institutions nearby (Card, 1993), since the availability of school in a locality can increase the level of schooling in general, because living far from education institutions increases the cost of schooling in different ways such as transportation cost, fatigues and homesickness;
5. bad habits such as smoking since according to the health economics, more educated people have better health and better health habits. Furthermore, Grossman (2008) asserts that completed years of formal schooling is the most important correlation of good health and this is based on some measurements, such as: mortality rates, morbidity rates, self-evaluated health status or psychological well-being.

Alternatively, the analysis can use policy or natural instruments such as: compulsory schooling law or other related education policies (Angrist and Krueger, 1991; Duflo, 2001; Comola and de Mello, 2010; Purnastuti *et al.*, 2015). The idea is that a child who was born earlier or before the policy is implemented, will have a lower level of education as compared to the people born later on.

Even when there are many alternatives of instrumental variables, searching for a valid instrument is hard. Weakly correlated instruments with the endogenous variable causes IV estimates to be biased in the same direction as the OLS and may not be consistent. Using many instrumental variables could also decrease the number of observations that has a serious effect for small sample studies particularly.

To evaluate whether the instruments used are appropriate, the standard quality, validity and relevance criteria of the instruments are considered (Purnastuti *et al.,* 2015). The most important thing is to use a valid instrument if it affects earnings through schooling only. The first test is for the quality of the instruments. This is assessed using an F-test of the joint significance of the respective instrument set in their first-stage equation. Moreover, the R squared from the first-stage equation for the IV models based on the conventional instruments must be at a reasonable level. In addition, test for overidentification restrictions and Hausman test for relevance criterion can be used. Hausman test can confirm the necessity to use IV/OLS estimations.

Sample Selection Issue

The problem arises due to non-random selection of the sample used for the estimation process, when the sample is only based on a subpopulation. For example: Mincer wage equation estimation is based on individuals whose wage is observed or individuals who choose to actively participate in the labour market as a wage earner. Thus, the differences between characteristics of actives and inactives may cause the sample selection bias. If the decision to or not to participate was a random decision, OLS estimates would be an appropriate estimating procedure; yet it is not a random decision, instead it is driven by some other factors (Bhatti, 2012). Moreover, Comola and de Mello (2010) assert that information on earnings is usually available only for salaried workers, thus OLS estimates are inconsistent if the earnings distribution is truncated.

Gronau (1973) firstly raised this issue. Most empirical studies are only based on the observed wage distribution. Meanwhile, the observed distribution represents only one part of the wage offer distribution, as the other part being rejected by the job seekers as unacceptable. As a result, the traditional estimation procedures may involve certain biases when applied to the secondary labour group such as married women, teenagers, and the aged. The study finds that the US females participating in the labour force were different in characteristics from those who decided not to be included in the labour force. Hence, simple OLS may produce biased estimates for different factors influencing wage in the labour market.

The corrective measures can be used in order to avoid the possibility of sample selection bias; the results may be considered for the whole target population. While ignoring this correction, it means that the result is valid only for the subpopulation of people who decided to work in the labour market. Heckman proposes maximum likelihood (ML) estimator (1976) and then a two-step estimation procedure (1979). The estimation procedure eliminates the possible sample selection bias in two steps. The first step uses ML probit regression, in which the decision to or not to work in the labour market is used as a response variable that depends on different explanatory factors. Then, the Inverse Mills Ratio (IMR) is calculated from the coefficient estimated in the first step. ML probit regression is estimated by separating the sample into two groups:

$$ACTIVE = \begin{Bmatrix} 1\ if\ person\ is\ involved\ in\ waged\ work\ in\ labour\ market \\ 0\ otherwise \end{Bmatrix}$$

Then, ACTIVE can be estimated by equation:

$$ACTIVE_{i,t} = \alpha_0 + \sum \beta_{1,n} K_{i,n,t} + \ \varepsilon_{i,t} \qquad (3.9),$$

where:

$ACTIVE_{i,t}$ is probit / dummy variable for active and not active,

$K_{i,n,t}$ is the number of explanatory variables that affect the likelihood of participation of individuals into waged work.

From the equation above, the Inverse Mills Ratio (IMR) is calculated by:

$$IMR_{i,t} = \frac{(V'\alpha)}{1-(V'\alpha)} \qquad (3.10),$$

where: (.) and (.) are density and distribution function of the standard normal distribution.

And the second step, estimating Mincer wage equation with the IMR as an additional regressor, or

$$\ln W_{i,t} = \beta_{0,t} + \sum \beta_{1,n,t} X_{i,n,t} + \sum \beta_{1,n,t} K_{i,n,t} \ + \beta_{3,t} IMR_{i,t} + \varepsilon_{i,t} \qquad (3.11),$$

this will account for the bias due to the non-random nature of the sample of wage earners. A significant coefficient for the IMR points at the presence of the sample selectivity. In short, a first-stage probit equation estimates the selection process, which is estimated as a probit (the dependent variable is binary variable for being selected into the sample or dummy variable is one if the woman is participating in the labour force and is thus being paid and observed for the wage equation). The results from the first equation are used to construct a variable that captures the selection effect in the wage equation. One example of instrument variables is the presence of children for women, since the assumption is women with kids may be more likely to be stay-at-home mothers; but working women with kids would not affect their hourly pay rate.

Wang (1995) studies the relationship between marital status, presence of children and women's wage rate in the US, using 1985-1992 National Longitudinal Survey of Youth's (NLSY) data. The sampling includes women aged 25 to 35 years old who exhibited no health limitations for job participation. Using the OLS and IV to control endogeneity and Heckman to control selection bias, the study finds that the marital status does not explain women's labour participation; women who have higher financial responsibilities such as having more children are more willing to work. By controlling the endogeneity of marital status and the number of children, the study finds that marriage does have a significant effect on wage as married women tend to have a lower market wage than the unmarried ones. According to Heckman correction model, the term is significant in the wage rate equation; thus, there is a difference for women who make the decision to or not to work; this implies that women who work either require higher wages or have higher ability level, or both.

Aslam (2009) combines some methods to estimate the rates of return on education in Pakistan and compares OLS with Two-Step Heckman method. The study controls the sample selectivity bias with some variables as exclusion restrictions, such as number of children aged less than 7 years old, number of adults aged minimum 60 years old in the household, marital status and natural logarithm of unlaboured income. The study shows that overall OLS overestimated the return to education compared to the other model and concludes that the return for females was consistently higher than that for males after controlling the bias from sample selection.

Dumauli (2015) addresses the sample selection issue to estimate the private return to education in Indonesia. The study uses IFLS4 data (year of 2007) and employs the Two-Step Heckman to control for the bias arisen from the non-random sample for female workers. In Two-Step Heckman, the first stage is to estimate the female labour force participation rate using a probit model to get a sample selectivity correction term. In this stage, the study adds some exclusion variables which influence the probability to participate in the labour market, but also which indirectly influence salary. The second stage is to include the IMR as an additional explanatory variable to estimate wage function. If the coefficient of the IMR is statistically significant, it means that the estimated results suffer from sample selectivity bias. The study uses four instruments, *i.e.*: natural logarithm of household assets, regional (provincial) unemployment rate, marital status and number of household members or household size. The reasons for choosing these variables are as follows: (1) higher household assets, which play as proxy for unearned (unlaboured) income, can discourage the possibility of females from participating in waged work; (2) the high unemployment rate increases the probability of females to participate in waged work; (3) number of household members and (4) marital status can influence the decision of females to join the labour market since these variables can influence females in terms of housework and the time that has to be allocated for their families.

Dumauli (2015) finds that the coefficients for natural logarithm of household assets, marital status and number of household members are negative and statistically significant, which means that these variables discourage females from participating in waged work, whereas the coefficient of (provincial) unemployment rate is positive, which means that this variable increases the probability for females to participate in waged work. The coefficient of the IMR is negative and statistically significant for both years of schooling and level of education. It suggests that the OLS' previous estimates suffer from sample selectivity bias. Comparing OLS with Two-Step Heckman, the coefficient of educational levels of the Two-Step Heckman is larger except for university level.

3.2.3 What is the Estimated Rate of Return?

There have been many studies which analysed the association between wage premium and higher education. As discussed above, the estimated rate of return (proxied by log of hourly wage) depends on education level and experience in the labour market, as explained in Mincer wage equation (Section 3.2.2). Thus, what will happen to wage is obvious; the higher the level of education, the higher the wage; even though the qualification is not in accordance with the required one, at least education acts as a signal that higher education qualification could imply higher competencies. One of the limitations of the present study is that it only considers education level and amount of wages, without considering the subject of study, education background and different quality of education institutions.

Psacharopoulos and Patrinos (2004) find that the private return to education varied in each region based on the latest data available in each country. The method used to estimate the return is the standard Mincer wage equation, following Psacharopoulos (1994). They define education variable as a dummy of education level, comprising the primary, secondary and higher education. By comparing adjacent dummy variable coefficient, the study could derive the private return to education. Psacharopoulos and Patrinos find that among the continents, the highest average private return to education was in the sub-Saharan Africa region, around 37.6 per cent for primary school, 24.6 per cent for secondary school and 27.8 per cent for university. Asia (non-OECD countries) had 20 per cent return to education for primary school, 15.8 per cent for secondary, and 18.2 per cent for university. In contrast, OECD countries (most of which are developed countries) have a lower return to education of only 13.4 per cent for primary school, 11.3 per cent for secondary and 11.6 per cent for higher education. It seems that return to education in OECD countries decreases as the level of education increases.

They also document private return to education in many countries, such as Australia which had return to education rate of 14.2 per cent for secondary school and 4.2 per cent for higher education in 1981. Returns to education in India in 1995 were 2.6 per cent for primary school, 17.6 per cent for secondary school and 18.2 per cent for higher education. Returns to education in Malaysia in 1978 were 32.6 per cent for secondary and 34.5 per cent for higher education. Unfortunately, the study does not have private return to education data for Indonesia, the US and the UK (Psacharopoulos, 1994).

For Indonesia, Psacharopoulos and Patrinos estimated the return to primary school for males was 19 per cent, and for female was 17 per cent in 1982; the return to secondary school for males was 23 per cent, and for females was 11 per cent in 1982; those returns significantly changed to 11 per cent for males and 16 per cent for females in 1986. And the return to university for males was 10 per cent, and for females was 9 per cent in 1982, with a slight change in 1986 to 9 per cent for males and 10 per cent for females. Those estimations were obtained from Psacharopoulos (1994).

The limitations of Psacharopoulos and Patrinos (2004) are: the data come from the firm-based sample in the hope to control the survey cost and the questionnaire is typically filled by the payroll department rather than by the individual employees. This approach could lead to the use of sample concentrated only in urban areas. Alternatively, if the rate of return is relative to no education, a person who has poor primary education would have their opportunity for getting worthwhile secondary education diminished, which could affect the rates of return.

In most countries, the majority of university graduates prefers employment in public sector. The concentration of graduates in public sector is identified as a problem in growth studies. However, public employee’s pay-based rate of return estimation is useful in private calculations. This is an incentive set by the state to invest in education and opt for employment in the public sector. Thus, the resulting contradiction is possibly due to these limitations.

Montenegro and Patrinos (2012) recalculated the return to education around the world, using the standard Mincer wage equation. The data used are from a large database constructed from existing national household surveys. It was prepared for the World Bank's World Development Report Unit over the period 2005-2011 and has been used in World Development Report during this period and also in several Human Development Reports for the period between 2000 and 2011. The sample is only waged employees, and the omitted variables of education is lower than primary school.

Based on education level, the study finds that the return to tertiary education on average in the world was around 16.8 percent, this is the highest one; meanwhile, the return to primary and secondary schools were 10.3 per cent and 6.9 per cent, respectively; which is in line with the human capital theory: wage increases with education level. Based on region, the return to primary school, secondary and tertiary education in East Asia were 11 per cent, 6.3 per cent, and 15.4 per cent, respectively. The remaining return in other regions is presented in Table 3.2.

Table 3.2: Return to Education by Level and Region (Latest Available Year between 2000-2011)

Region	Primary	Secondary	Tertiary	GDP/pc (PPP 2005)	N
World	10.3	6.9	16.8	6,719	74
Middle East and North Africa	9.4	3.5	8.9	3,645	7
South Asia	9.6	6.3	18.4	2,626	4
Eastern and Central Europe	8.3	4.0	10.1	6,630	7
High Income Economies	4.8	5.3	11.0	31,748	6
East Asia and Pacific	11.0	6.3	15.4	5,980	6
Latin America and Caribbean	9.3	6.6	17.6	7,269	20
Sub-Saharan Africa	13.4	10.8	21.9	2,531	24

Source: Montenegro and Patrinos (2012).

In general, return to secondary education is the lowest one, except in high income economies. The result would imply that the high return in tertiary education is because of scarce supply; based on continents, the highest returns are in Sub-Saharan African and South Asia. This finding supports Psacharopoulos and Patrinos (2004) as well, that the returns to education in high income countries tend to be lower than in developing countries The possible reason for this is workers with high education attainment is relatively more available in developed countries; thus, the returns to human and physical capital tend to be equated at the margin. Another issue is measuring different quality level of education with the same model of the Mincer wage equation leads to biased results. As Glewwe (1996) argues, education quality in developed and less developed countries is low and uneven in quality. Thus, the omission of the school quality variable in Mincer wage equation leads to biased results. This could also invoke the errors in variables argument; the years of schooling (education) variable measures human capital attainment with a large amount of error, biasing the parameter estimates on years of schooling towards zero (the tendency of unmeasured ability to cause the opposite bias may have been prevented by controlling for sample selectivity).

Moreover, several India's studies are also discussed in this part as a comparison, because the characteristics of India and Indonesia are similar, for example: both countries are in Asia, have abundant population, and transform the economy from agriculture to the service sectors. According to the World Development Indicators, India's share in agriculture, manufacturing and services were 17 percent, 30 percent, and 53 percent in 2015, respectively, whereas Indonesia's share in agriculture, manufacturing and services were 14 percent, 40 percent and 46 percent in the same year, respectively.

Rani (2014) and Singhari and Madheswaran (2016) study the return to education in India. Specifically, Rani (2014) looks into the impact of different levels of education, religion, caste as well as the impact of living in urban and rural communities on earnings in India, using a large cross-section sample of India Human Development Survey to estimate the Mincer and the augmented Mincer equations. The study finds that the return to education of primary school (completed years of education between 1 and 8 years) was around 1.3 percent, return for secondary school was 3.7 percent, and return for graduate or university was 15.4 percent. This result confirms that the correlation between education and wages are positive.

Singhari and Madheswaran (2016) estimate the standard Mincer wage equations separately for regular and casual workers. The result is slightly different with Rani's (2014): the return for primary school is relatively higher; however, the trend decreases. For regular workers, the highest rate of return to education is diploma, followed by graduate and post-graduate, and then secondary education, as shown in Table 3.3. On the other hand, other studies' estimations are almost the same with those of Singhari and Madheswaran (2016).

Table 3.3: The Estimates of Return to Education in India Based on Some Studies

Author (s)	Primary	Middle	Secondary	Graduate
Blaug (1972)	16.5	14.0	10.4	8.7
Psacharopoulos (1973)*	24.7	19.2	-	14.3
Husain (1967)	-	-	4.8	12.0
Tilak (1987) unadjusted estimates	33.4	25.0	19.8	13.2
Tilak (1987) adjusted estimates	7.8	8.5	Negative**	6.8
Recent studies				
Duraisamy (2002) unadjusted estimates	7.9	7.4	17.3	11.7
Duraisamy (2002) adjusted estimates	7.8	7.4	17.7	12.7
Agrawal (2011)	5.5	6.1	12.2	15.9
Rani (2014)	1.3[a]	-	3.7[b]	15.4[c]
Author's calculation (Regular Workers) Standard Mincerian 1983	11.5	4.0	7.1[d]	10.0
Author's calculation 1993-94	8.1	3.6	9.3	7.1
Author's calculation 2004-05	15.1	5.2	12.0	12.6
Author's calculation 2011-12	9.3	5.2	10.2	11.5

Source: Rani (2014)

Notes: Rani (2014) has taken log hourly wage as dependent variable in wage equation.

* As quoted in World Bank Staff Working Paper, 1979, No. 327.

** Tilak (1987) did not report the actual returns.

[a] elementary education i.e. completed years of education between 1 and 8 years.

[b] secondary education i.e. completed years of education between 9 and 12 years.

[c] Higher education i.e. completed years of education with 13 years and above.

[d] Secondary education includes higher secondary education.

Source: Rani (2014) in Singhari and Madheswaran (2016).

With regards to the empirical studies of return to education in Indonesia, Comola and de Mello (2012) study the wages in Indonesia by considering the selection bias. The study uses the 2004 wave data from the Indonesian National Labour Force Survey (SAKERNAS). To deal with the endogeneity of educational attainment, the study instrumented years of education by exposure to primary school construction program (Sekolah Dasar INPRES), measured as the intensity of school construction in an individual's district of birth and his/her age when the programme was launched. Thus, this programme affected the improvement of school attainment as well as the increase of wages, as explained in Section 3.2.

The study does not specifically find the amount of the wages or return to education, instead it compares the results of several methods, *i.e.* multinomial selection and binomial procedure (Heckman, 1979). The study finds that several parameter estimates differ when multinomial selection is allowed in the estimation of the wage equation. Moreover, one of the most important findings is that the estimated returns to education do not seem to be affected by the selection bias; it is very interesting for future empirical research to identify any strategies that can deal with the endogeneity of education attainment in wages equations independently from the complexity of underlying selection process, if any.

Furthermore, Dumauli (2015) estimates the private return to education in Indonesia by taking into account the endogeneity and sample selection issues, thus using methods other than OLS for Mincer wage equation; they use parental education as the instrument variable in IV method; and household size, marital status, and regional unemployment rate as an instrument of selection bias in the Heckman method. The data used are the fourth wave of the IFLS (IFLS4) year of 2007. The study finds that the return to education was 11.7 per cent in general (OLS), which means each additional year of schooling increases wages by 11.7 per cent. In terms of education level, the study finds that the return to primary school was 26.6 per cent, return to junior high school was 64.2 per cent, return to senior high school is 109.2 per cent, and return to university was 150 per cent higher than individuals with no schooling. Meanwhile, Heckman's result is slightly lower: 11 per cent; and 17.3 per cent for IV method. However, parental education does not pass the over-identification test, in other words, the instrument is not appropriate.

Dumauli's (2015) finding has an implication that the higher the level of education is, the higher the rate of return to education becomes, thus agreeing with the human capital theory. That being said, compared to previous studies or to other countries within the region, the return to education in Indonesia was still relatively low. Dumauli (2015) argues there are two potential explanations for the low return to education in Indonesia. Firstly, the low rate of return to education in Indonesia may be induced by the low quality of the education system. Another possibility is that high skilled jobs are not in demand in Indonesia, but rather exist in the supply side. In particular, there is an excess supply of college graduates, and this in turn reduces the wages of college graduates.

In terms of gender difference, Dougherty (2005) documents previous empirical studies with male and female schooling coefficients. The study finds that the estimates of the return of schooling in the US tend to be higher for females than for males; for example: Angle and Wissman (1981) estimate Mincer's equation with hourly wage as dependent variable and years of schooling/level of education, age, family background and ethnicity as the independent variables. The data used are the NLS young men and young women with 2,831 males and 1,677 females who had at least some college qualifications. The study finds that the coefficient of years of schooling for male was 0.040 and for females was 0.076. In addition, the coefficients for females with bachelor and master degrees are higher than males' coefficients, except for the coefficient of females with a PhD. However, Abbasa and Foreman-Peck (2008) argue that high measured returns to female education is caused by a combination of much lower workforce participation and fewer average years in the labour force than males.

Card (1999) uses CPS March data in the US from 1994-1996 period. The dependent variable is hourly wage, the explanatory variables are years of schooling, potential working experience, squared of experience, and ethnicity. The study concludes that the coefficients of years of schooling were 0.100 for males and 0.109 for females.

Similar to the US findings, Trostel *et al.* (2002) estimate the returns to schooling in 28 countries (mostly European), with data derived from a common survey instrument. They find that the return to education of females is higher than males: females' return is 5.7 per cent, while males' return was 4.8 per cent. Psacharopoulos and Patrinos (2002) list 95 estimates of males and females schooling coefficients from 49 countries at different dates. Of these, 63 were greater for females, 3 were equal, and 23 were greater for males.

Dougherty (2005) asserts that there are some possible causes of this effect (coefficient of return to education for females tends to be higher than for males):

1. sample selection bias, since the common development in literature is to treat the wage equation as the second stage of a two-stage model of labour force participation and earnings. If labour force participants differ from non-participants, OLS estimates are likely to be inconsistent. Given that most males do participate, while many females do not, the strength of the bias could be different between the genders and this might be a factor responsible for part of the difference in the OLS coefficients;
2. schooling and discrimination; a second extension of the standard Mincer equation is to allow for the possibility that schooling may have two effects on earnings, at least for females: a direct human capital effect, and an indirect, anti-discrimination effect. There are two possible reasons for this; firstly, the better educated the individual is, the more likely the individual is to have a degree or formal qualification that would help to standardise wage offers, regardless of sex. Secondly, the better educated a female is, the less likely she is to be tolerant of discrimination. Moreover, the better educated a female is and the greater her potential earnings are, the more able she is to pay for childcare and other services that allow her to seek a wage offer that fully values her characteristics.

3. quality of the schooling investment; there may be a difference in the quality of male and female education. It is assumed that females tend to be more motivated students than males and extract more from their time in school. As such, measuring schooling in terms of years of enrolment may mask systematic differentials in the quality of the schooling investment, and if the quality of the investment correlates with years of enrolment, its omission from the regression specification could cause differential biases in the male and female schooling coefficients;
4. occupational choice; females may be under-represented in jobs where schooling is a relatively unimportant factor in the determination of earnings. For example, they may be under-represented among union workers, where schooling is subordinated to seniority as a determinant of earnings, or in self-employment where entrepreneurial skills are relatively more highly valued;
5. the other possibility is the differentials in measurement error and endogeneity of schooling and work experience. However, these possibilities are pursued in the study.
6. meanwhile, Ren and Miller (2012) investigate the gender differential in the average return to schooling in China, and argue that the difference occurs possibly because of the differences in the demand for, and supply of, education between males and females, greater positive self-selection of females into the labour force relative to males, a more limited supply of skilled female workers, different technological requirements between the female-dominated and male-dominated jobs, and discrimination against female workers that is less intense among the better educated.

Abbas and Foreman-Peck (2008) also study the implication of education policy in Pakistan by the Mincer equation and using Pakistan Social and Living Standards Measurement (PSLM) Survey, between 2004 and 2005. In terms of gender, they conclude that returns to an additional year of schooling are greater for females than for males in a paid employment; the returns to an additional year of schooling were 9.2 percent and 14 percent for males and females, respectively. The large and significant gender differences in Pakistan cannot be attributed to the scarcity of educated women; a possible explanation is the low probability or short duration of employment which requires a high apparent rate of return compared to males. Unfortunately, Abbasa and Foreman-Peck (2007) did not have ethnicity as well as religion as control variables, they have Punjab, Sinndh, NWFP, and Balochistan as provinces variables/dummies. Abbas and Foreman-Peck (2008) also conclude that there is some signalling in Pakistani education investment but mainly the education is a productivity-enhancing investment in human capital. The conclusion is based on the comparison of coefficients of schooling estimated for the paid employed with those for the self-employed. The result shows that the coefficients of schooling for the self-employed are lower than for the paid employed. Thus, there is apparently a small element of signalling. And the bulk of the return to schooling is a consequence of enhanced human capital.

In addition, the great recession may affect the return to education to some extent, in particular wage gap by education level. Belfield (2015) estimates the return to education in Arkansas over the period before, during and after the great recession. Belfield reveals very modest effects of the great recession on the earnings gaps of workers with different level of education. During 2001-2012 period, there were large and stable returns to postsecondary education relative to high school completion, and these gaps were largely unaffected by the Great Recession. The study also found that employment shocks that differ by education level: for persons without a college education, employment shocks were stronger, and they persisted beyond the end of the recession. Adjusting for these employment shocks, earnings gaps by education level increased over the period after 2007. Moreover, there is an evidence that those who graduated from college during the Great Recession gained less than those who graduated before 2007. As with earlier recessions, postsecondary education served as an effective buffer against labour market shocks.

3.2.4 What is the Likely Impact of the Expansion?

The Simple Supply-Demand Framework

Over the last decade, the education sector has been expanding both in developed and developing countries including Indonesia and this expansion has clearly affected the labour market. One of the alternatives to analyse the effects is the simple supply and demand framework in a competitive market. The framework has been widely used in academic research on changing labour market inequality (Machin and McNally, 2007). The assumption is that competitive market could be too strong for the Indonesian labour market, considering Indonesia is one of the small open economies in the world. Nevertheless, International Business Publication (2017) asserts that the Indonesian labour market is generally open and flexible, although there are significant restrictions on the use of contract workers.

The present study uses examples of an expansion in university graduates relative to primary school (or below) graduates (relative employment) and its effect to relative wages in this part. Beginning with a position where the demand and supply are perfectly equalised or the initial equilibrium condition; an increase in the supply of university graduate leads to a decrease in the wage premium, *ceteris paribus*. This is because employers have a wider range of similarly qualified individuals to choose from (Machin and McNally, 2007). The wages of university graduates (Uni) and primary school (or below) graduates (Pri) are denoted by $W_{university}$ and $W_{primary}$, and the employment rates are l_{uni} and l_{pri}, respectively. The initial equilibrium (E_0) occurs at the intersection of the initial relative demand (D_0) and relative supply (S_0) curve, with associated relatives wages $(W_{uni}/W_{pri})_0$ and the relative employment $(l_{uni}/l_{pri})_0$, as shown in Figure 3.3.

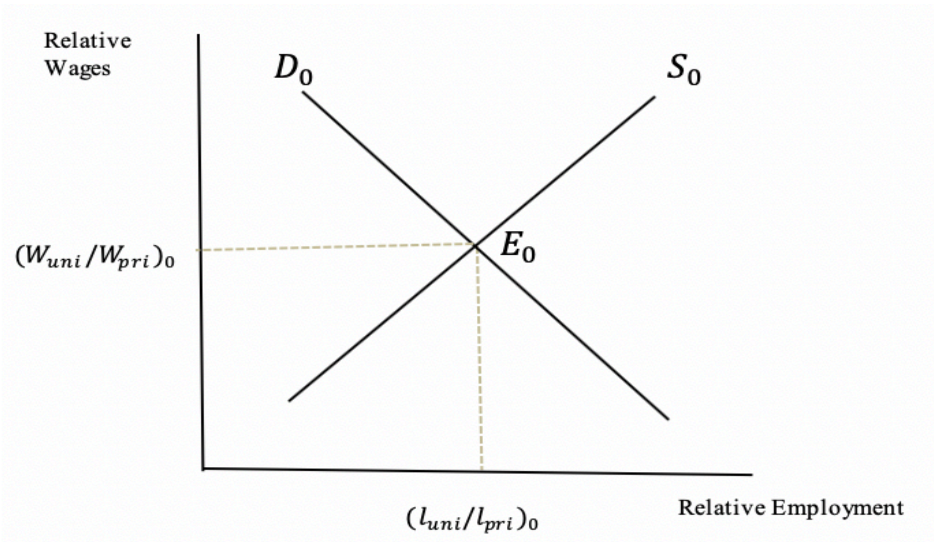


Figure 3.3: The Initial Condition in the Labour Market

When there is an expansion in university level, it is followed by an increase in the supply of employees with university qualifications. Subsequently, the supply curve shifts to the right or from S_0 to S_1. The relative employment rate increases from $(l_{uni}/l_{pri})_0$ to $(l_{uni}/l_{pri})_1$, while the relative wage decreases from $(W_{uni}/W_{pri})_0$ to $(W_{uni}/W_{pri})_1$, and the new equilibrium is E_1. In short, this model predicts that the supply shock reduces the relative wage of university graduates, as shown in Figure 3.4.

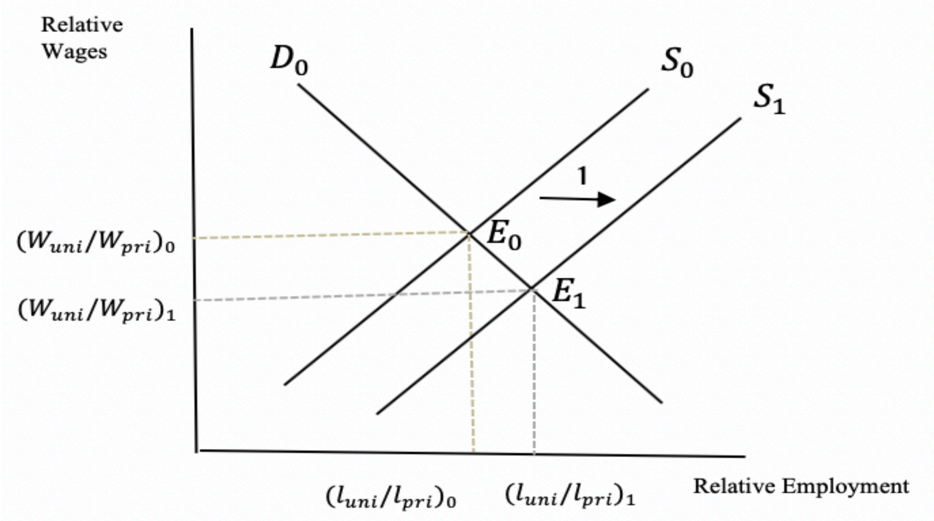


Figure 3.4: The Supply Curve Shifts Due to the Higher Education Expansion

Following the shift on the supply side, employers tend to demand more employees with higher education qualifications. Then, there is an increase in the relative demand and the demand curve shifts to the right from D_0 to D_1, as shown in Figure 3.5; as such, lower relative wages will not be the case, assuming that the relative supply curve is fixed (S_1). If the demand increases faster than the supply, the wage premium will increase, and the new equilibrium occurs in E_2 with higher relative wages $(W_{uni}/W_{pri})_2$ but the same employment rates $(l_{uni}/l_{pri})_2$.

It is worth noting that the elasticity of the relative demand and the relative supply of higher education graduates would have significant effect on the analysis. The elasticity refers to the degree of responsiveness in supply or demand in relation to changes in wage premium. If the relative supply or demand curve is more elastic, then small changes in the relative wages premium will cause large changes in the number of relative employment. *Vice versa,* if the relative supply or demand curve is less elastic, then it will take large changes in the relative wages to cause a change in the number of relative employments. Moreover, the elasticity of wage premium with respect to demand and supply is not necessarily the same across countries and over different time periods.

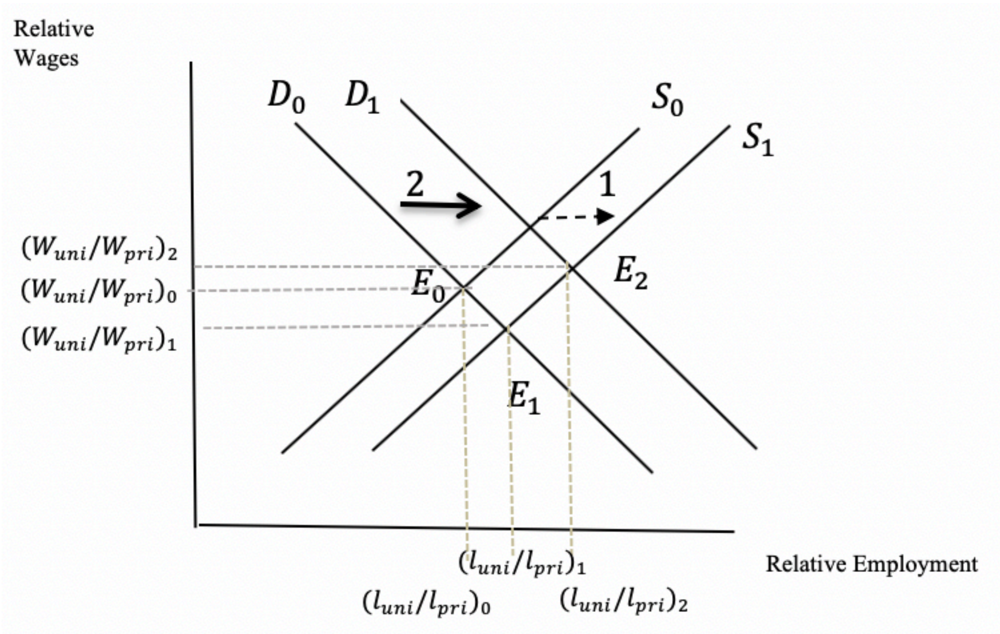


Figure 3.5: The Demand Curve Shifts Following an Increase in the Supply Side

3.2.5 What has Happened to the Rate of Return after the Expansion?

The simple supply and demand framework specifies that if the relative demand increases faster than the relative supply, then the return to education (wages) will increase, as explained in Section 3.2.4. This part, then, will discuss what has happened to the rate of return after the expansion of education by taking into account empirical studies in the US, the UK, and OECD countries, followed by empirical studies from Indonesia and other similar countries.

Many studies have analysed this expansion and some of them even prove that wages have increased after the education expansion (Katz and Murphy (1992), and Machin and McNally (2007)). On the other hand, other studies also find that wages have decreased (Walker and Zhu, 2008).

Firstly, Katz and Murphy (1992) use a simple supply and demand framework to analyse the change in wage structure in the US from 1963 to 1987. The data used were a series of 25 consecutive March Current Population Surveys (CPSs). The education variable used was years of schooling which is then divided into 4 groups: (1) 8-11 years for dropout group, (2) 12 years for high school graduate, (3) 13-15 years for some college, and (4) 16+ year for university graduate; while the wage is measured by the average, weekly wage of full-time workers within gender-education-experience. The sample includes all individuals who worked at least one week in the preceding year. The study finds the college wage premium increased moderately in the 1960s, decreased in the 1970s and increased significantly in the 1980s. The driving force behind the observed change in wage structure is the rapid secular growth in the demand for more-educated workers as well as female and more skilled workers. In addition, they estimate that the elasticity of substitution between skilled and unskilled labour was significant, around 1.4 – 1.6, which implies that the increased supply reduces relative wages (*ceteris paribus*). Thus, the differences in the growth rate of the supply of university graduates have an important role to play in explaining the changes in college premium in the US during that period. The demand has outstripped supply; hence, the wage premium increased significantly over this time period despite the large increase in supply.

Machin and McNally (2007) have addressed some issues related to the expansion in tertiary level education: over-supply, over-qualification and skill mismatch. Machin and McNally also estimate the change of returns to education in OECD countries in their study, using the simple relative supply and demand framework. Moreover, they provide empirical evidence for each country. In general, they find that the wage premium has increased in most OECD countries, except for Spain, New Zealand, and South Korea. In South Korea, the decline of wage premium occurred during rapid industrialisation period (1974-1990). Yet, there is still a positive return to tertiary education, even in countries with a decline in return. The strong, positive and increasing returns to higher education suggest that 'under-supply' is more of an issue and that continued expansion is justified.

For the UK, Machin and McNally (2007) document previous research and assert that the proportion of employees with higher education qualifications in the UK increased from 5 per cent to 21 per cent over the period of 1980-2004. The relative weekly wages also increased from 1.48 in 1980 to 1.64 in 2004 in the UK. This concludes that the pattern of change in the wage structure differs; however, the pattern of change in university wage differentials is fairly clear. There was a significant increase in the 1980s, in terms of the average wage return, which continued to increase at a lower rate in the 1990s; and became relatively stagnant in the 2000s.

Some more debates have emerged in more recent works about certain subgroups and whether there is any evidence of falling returns to education. A number of studies documents rising returns over time from the 1970s to the early 1990s such as: Harkness and Machin (1999). A number of other studies also find slightly rising or constant returns from the early 1990s to the early 2000s, such as: Chevalier *et al.* (2004) and Walker and Zhu (2003). In contrast, there is little evidence of declines in return to education such as: O'Leary and Sloane (2005) and Walker and Zhu (2005), even Dickerson (2005) reports no change in return to education using the same data sources. Moreover, reports of falling returns need to be kept in perspective as the size of returns are still substantial in comparison to those with only an upper secondary education. Overall, the UK pattern of change in relative wage and employment was similar to the US in that the relative demand for higher education has outstripped the relative supply.

Particularly for South Korea, Kim and Topel (1995) study the development of wage premium during Korea's rapid industrialisation, from 1970 to 1990. The analysis is based on the simple supply and demand framework, whilst ignoring capital and assuming the production in industry is homothetic in its labour input. The study find that there were; a very substantial upgrading of skills in the workforce, an increase in the number of university graduates, a positive wage premium, and an increase in the wage premium at the end of that period. Kim and Topel argued that the change in wage inequality can be driven by (at least) three factors, *i.e.* changes in relative supplies of different skill groups can change their relative wages; expanding industries may be intense users of some employee types which in turn raises the relative demand for those employees; and finally, the pace of technical change may favour certain employee types. As such, there is a shift in relative labour demands over time.

With regards to Indonesia and other East Asian countries, Gropello and Sakellariou (2010) estimate skill and wage premium at the national and sectoral levels in seven East Asian countries: Indonesia, the Philippines, Vietnam, Cambodia, Thailand, China and Mongolia. For Indonesia, the study used a 2-digit industry classification, and the data used are from SAKERNAS surveys from 1994 to 2007. Slightly differently, the method used is the linear regression with employees' characteristics (including age and gender), type of labour (skilled or unskilled), industry indicators (dummies) and the social return to education (skill premium) as the independent variables; and wage premium as the dependent variable. The study confirms that there is a decrease of workers with primary and below primary school qualifications. In contrast, there is an increase (expansion) in workers with high school and university qualifications.

Gropello and Sakellariou (2010) findings are summarised in Tables 3.4, the base (reference) education level is completed primary or lower education level; except that the presented return estimates for education levels higher than primary are given in comparison to primary. Returns to education decreased in most education levels, with the largest decline during 1994-2007 period being the junior and senior high schools (relative to primary school). Meanwhile, the lowest decline was the university level. In addition, the study separates the sample within senior education and finds that most declines in premiums were associated with the vocational-technical education. One possible reason for this is that the demand for higher education workers is on the rise but only moderately, or that the supply of labour outstripped the demand for labour.

For other Asian countries, the Philippines had an increasing trend for senior and tertiary education in basic regression but turning negative when the study adds a control variable and industry dummies. It also seems that Thailand had similar trend to Indonesia, as explained previously, a negative growth in return to education for most of education levels. In contrast, Vietnam and Cambodia had a positive trend or an increase return to education in all education levels. Finally, China had a positive return as well, except for the primary school qualifications, as shown in Table 3.5. Gropello and Sakellariou (2010) argue that those evidence combined with stable or increasing education/skill wage premiums (in regressions with only basic controls) indicate a generally increasing demand for skills in the region (and that education also leads to increasing inequalities in several countries).

Table 3.4: Return to Education by Education Level (Relative to Primary Education) in 1994, 2001 and 2007

	1994	2007	Change (%) 1994-2007
Basic controls only			
Primary	0.213	0.242	13.6
Junior/Primary	0.234	0.128	-45.3
Senior/Primary	0.596	0.508	-14.8
University/Primary	1.125	1.084	-3.6
Basic controls + Industry dummies			
Primary	0.195	0.243	25.1
Junior/Primary	0.235	0.126	-46.4
Senior/Primary	0.61	0.498	-18.4
University/Primary	1.137	1.07	-5.9

Basic controls only

High school general/primary	0.527	0.486	-7.8
High school vocational/primary	0.663	0.547	-17.5
University/Primary	1.123	1.084	-3.5
Basic controls + Industry dummies			
High school general/primary	0.545	0.479	-12.1
High school vocational/primary	0.678	0.526	-22.4
University/Primary	1.139	1.056	-7.3

Source: Gropello and Sakellariou, 2010.
Note: Return to primary school relative to below primary school graduates.

Table 3.5: The Change of Return to Education by Education Level in Several Asian Countries

Education Level	Indonesia	The Philippines	Thailand	Vietnam	Cambodia	China
	Change					
	1994-2007	1998-2006	1990-2004	1999-2006	1997-2007	1999-2005
Basic control only						
Primary	0.029		0.093	0.113	0.154	0.036
Junior/Primary	-0.106		-0.058	0.227	0.135	0.105
Senior/Primary	-0.088	0.185	-0.033	0.385	0.277	0.274
University/Primary	-0.041	0.101	0.196	0.648	0.722	0.351
Basic control + industry dummies						
Primary	0.049		0.032	0.098	0.038	-0.019
Junior/Primary	-0.109		-0.092	0.206	0.056	0.111
Senior/Primary	-0.112	-0.209	-0.141	0.318	0.061	0.290
University/Primary	-0.067	-0.141	-0.017	0.487	0.411	0.392

Source: Gropello and Sakellariou, 2010.

Similar to Gropello and Sakellariou's (2010) conclusion, Purnastuti *et al*. (2013) confirm that there was a decline in the return to education in Indonesia. The estimation uses the standard Mincer equation with log of monthly earnings as dependent variables; education variable is in dummy of education level from primary to master level (relative to below primary school) and the control variables are: experience, marital status, gender and residential, by comparing IFLS1 (year of 1993) and IFLS4 (year of 2007) data. In addition, the study also uses the cohort effect since changes in the pay-off to education over time can reflect both cohort and age effects. The study finds that the returns to education tend to increase as the level of education increases. Also, gender affects the difference in return to education. Returns to education appear to be a less profitable investment in 2007 than returns in 1993, except for university degrees, but the profitability of this education level increased between 1993 and 2007, for both males and females. The study argues that the decreasing trend could be attributable to the large-scale expansion of education in Indonesia, or to a rate of expansion in the number of jobs requiring higher educational attainment which lagged behind the expansion of education and the increase of average education attainment.

3.3 Method and Data

3.3.1 Method

The present study applies two specifications of the Mincer wage equation, the first model is the standard Mincer equation (Equation 3.6), as explained in Section 3.2.2 and the second one is with control variables, including personal characteristics (P), work related and firm size (F), and regional dummy variables (A), that is formulated by:

$$\ln W_{i,t} = \beta_{0,t} + \sum \beta_{1,n,t} X_{i,n,t} + \beta_{2,t} L_{i,t} + \beta_{3,t} L_{i,t}^2 + \sum \beta_{4,n,t} \mathrm{P}_{i,n,t} + \sum \beta_{5,n,t} \mathrm{F}_{i,n,t} + \sum \beta_{6,n,t} \mathrm{A}_{i,n,t} + \varepsilon_{i,t} \quad (3.12),$$

where: $lnW_{i,t}$ is log of real hourly wages; Σ represents set (vector) of explanatory variables; $X_{i,n,t}$: education level of individual i at time t, n is 1...4 (primary to university level) or years of schooling (n = 1), $L_{i,t}$ is potential labour market experience in terms of year. Furthermore, other control variables are divided into three groups: personal characteristics ($\sum \mathrm{P}_{i,n,t}$), work related and firm size ($\sum \mathrm{F}_{i,n,t}$); and area/regional dummies ($\sum \mathrm{A}_{i,n,t}$), i is individual (1…I); and t is at time t (2000 and 2014). The explanation for each variable as follows:

Dependent Variable ($lnw_{i,t}$)

The dependent variable in this study is the natural logarithm of hourly wages in the real term (2010=100). In the estimation tables, this variable is notated by log of real hourly wage, for easier reading and interpretation. Respondents of both IFLS waves are explicitly asked about their last year wages (includes all benefits), total number working week per year, and working hours in a week. Hourly wages[14] are defined based on the formula:

$$Hourly\ wage = \frac{Last\ year\ salary}{Total\ number\ \ hour\ per\ week * Total\ number\ week\ per\ year}$$

(3.13),

$$Real\ hourly\ wage = \frac{100}{inflation} * Hourly\ wage,$$

inflation is in terms of index (2010=100), and wage is measured in Indonesian Rupiah (IDR).

The present study does not consider using monthly or yearly wages like most existing studies in Indonesia since hourly wage is a better measure of the return to education, as elaborated in Section 3.2.2. As wages are influenced by both wage rate and hours worked, hourly wage could eliminate unobserved heterogeneity caused by the omitted work hours (Li and Urmanbetova, 2007).

Main Independent Variables

Education Attainment ($\sum X_{i,n,t}$)

Education attainment is defined based on question dl16 and is divided into four main groups:

1. primary school and below, consisting of primary school, no schooling, adult education A, Islamic elementary school, and kindergarten,
2. junior high school, consisting of general junior high school, vocational junior high school, adult education B, and Islamic junior high school,
3. senior high school, comprising general senior high school, vocational senior high school, adult education C, Islamic senior high school, and Islamic school (pesantren),
4. university, consisting of college (diploma1, diploma2, diploma3), university (undergraduate degree), university (master's degree), university (PhD), and open university.

[14] Last year salary is defined by question TK25A; total number of hours per week is defined by question TK22A; and total number of weeks per year is defined by question TK23A.

Furthermore, other education levels and schools for the disabled are defined as missing value, since the disabled and general schools have different characteristics and the return to education will differ. Education level is used in the main analysis for simplification as explained in the introduction; this allows non-linear effects of the level of education (Steiner and Wagner, 1996) and is also in line with the policy implementation which is commonly based on the education levels. For simplification purpose, the present study also adds the analysis of years of schooling especially for robustness test. In the estimation tables, these variables are notated as: primary school, junior high school, senior high school, and university; or years of schooling.

Experience and Experience Squared ($L_{i,t}$ and $L_{i,t}^2$)

The standard approach to control Mincer equation is experience and experience squared. However, it is impossible to get complete labour market experience from both the 2000 and 2014 surveys. As such, potential experience is used in the regression (as explained in the literature review). The present study also follows Dong's (2016) definition, *i.e.* experience of workers with primary school degree or below is age (in year) – 12 years; experience of workers with junior high school degree is age – 15 years; experience of workers with senior high school degree is age – 18 years; and experience of university qualification is age – 22 years. Then, the square term is employed to construct experience squared, notated as L and L^2 in the estimation tables as experience and experience square.

Other Control Variables

Control variables are defined based on the literature and previous empirical studies. The present study also considers that those variables must be present in both IFLS surveys for comparative purpose. The study then categorises the control variables into three groups:

1. personal characteristics, consisting of dummy variables for: sex (1=female), and marital status (1=single, 2=married and cohabitate, 3=other status) as per Comola and de Mello (2009), Purnastuti *et al.* (2013), and Dumauli (2015); religion dummies (1=Islam, 2=Christian/Protestant, 3=Catholic, 4=Hindu/the omitted, 5=Buddhist, and 6=Konghucu), and Ethnicity dummies (1=Jawa, 2=Sunda, 3=Batak, 4=Betawi, 5=Minang, 7=Tiong Hoa, and 8=other/the omitted dummy) following Suryadarma *et al.* (2006) and Patrinos (2016). Ethnicity dummies are

defined based on majority ethnicities, except for *Tiong Hoa* or Chinese-Indonesian ethnicity which is only less than 1 per cent of the population, though the ethnicity has a special characteristic due to the argument that 70 per cent of the Indonesian economy is in the hands of the Chinese and that negative views towards the ethnicity proliferate again recently (Herlijanto, 2016).

2. work related and firm size variable consists of full-time/part-time dummy (1=full time[15] and 0=part time) as per Muffels and Fouarge (2001), tenure or current experience in the labour market (Purnastuti *et al.*, 2013), sector dummy (1=private and 0=public), industry dummies (agriculture; mining and quarrying; manufacturing; electricity gas and water; construction; wholesale retail restaurant and hotels; transportation storage and communications; finance insurance real estate and business services; and social services; with agriculture and other sectors as omitted variable), following Gropello *et al.* (2011); and firm size refers to the number of workers in the firm[16] (Ekberg and Salabasis, 2001);
3. regional dummy variable consists of urban/rural dummy or notated as Urban (as per Purnastuti *et al.* (2013) and Dumauli (2015)) and province dummies (Patrinos (2016) and Harmon *et al.* (2000)). However, the present study has a slightly different group of provinces between 2000 (IFLS3) and 2014 (IFLS5), though it consistently uses Nusa Tenggara Barat (NTB) as the omitted variable.

Furthermore, the present study employs several different estimation specifications, with only the main variable and personal characteristic controls; and both of them are in addition to work related and firm size. The result will not be discussed in this paper, but is attached instead in the appendix for further information (Appendix VI). Full regression is chosen with region dummies, since the result is better, considering consistency for the interpretations; and the study can analyse different effects of those variables on the wages. Moreover, this study also separates the estimations into gender and sector to elaborate the gender difference, to get more information on private and public sectors in Indonesia, and to see whether those sectors follow market mechanism in wage determination.

---

[15] Full-time workers are defined as working equal to or more than 30 hours per week; this is based on ILO which use 30 hours per week as the cut-off point for its definition of a part-time worker (Felipe and Hasan, 2006).

[16] The present study classifies into four groups: firm with 1-4 workers is an omitted variable, firms with 5-19 workers (small enterprise), 20-99 workers (medium) and more than 100 workers (big enterprise).

3.3.2 Data and Sample Restriction

IFLS3 (2000) and IFLS5 (2014) are the main data of the present study. Both surveys were fielded in 2000 and in late 2014 to early 2015, respectively. These waves are used to examine the effect of the National Education System Law (No. 20 of 2003) and Higher Education Law (No. 12 of 2012) enforcement. Thus, the year 2000 represents the period before the law was enforced and 2014 represents the years after the law was enforced, as explained in Section 3.1.

The main variables of the present study are wage, education attainment, and experience; this is based on the Mincer wage equation as explained in Section 3.2.2. The study also puts some restrictions on the data, and determines certain definition of employment in the research, since it focuses on the return to education from employment in private and public sectors in Indonesia. The full sample data (individuals who provided their education attainment) are provided in Appendix V.

Firstly: age restriction, this study restricts the sample to workers aged 16 - 55 or adult respondents. 16 years old is the minimum age of the sample, considering that people finish junior high school (as compulsory education) in 16 years of age, and 55 years old is the maximum age limit; this is based on Law no. 3 of 1992 on Social Security and Law no. 11 of 1992 on Pension Fund[17]. Workers whose age under 16 or older than 55 are exclude in the first restriction; thus, the sample is only workers in the labour market.

Secondly, the sample is restricted only for employment, following Dong (2015) who uses question TK01-TK04[18] in each IFLS wave. The restrictions ensure the sample is in the labour market or employed during the period. The third restriction is based on sector (question TK24A): "Which category best describes the work that you do?"[19] The present

[17] Those regulations are amended by Government Regulation no. 45 of 2015 article 15 on Pension Security. The regulation stipulates: "pension age is 56 years old in 2015; however, starting on 1 January 2019, the retirement age will be 57 years. This retirement age will gradually increase by one more year every three years until the retirement age finally becomes 65."

[18] Question TK01: "What was your primary activity during the past week?"; TK02: "Did you work for at least 1 hour during the past week?"; TK03: "Do you have a job/business, but were temporarily not working during the past week?"; and TK04: "Did you work at a family-owned (farm or non-farm) business during the past week?" And employments are respondents who only answer working/trying to work/help to earn income in the first question; and respondents who answered yes (working) for the remaining questions.

[19] Answers could be: self-employment, self-employment with unpaid family worker/temporary worker, self-employment with permanent worker, government worker, private worker, casual worker in agriculture, casual worker not in agriculture, and unpaid family worker. The answer choices are slightly different with those of IFLS3 which has no choice of casual workers; only self-employment, unpaid workers, government and private workers. However, the present study can ignore this difference, since this will not affect the following analysis.

study focuses on government (public) and private sectors in the waged sector, since they have a similar characteristic of receiving wages regularly, for instance monthly wages. Yet, it is difficult to distinguish between formal and informal sector among those firms because the other information is limited, there is a contract question/data, but not all respondents who answer that question, as a consequence the sample size may shrink. And a small sample size also affects the reliability of a survey's results because it leads to a higher variability, which may lead to bias. The final sample in the present study is around 90 per cent of employment in 2000 and 85 per cent of employment in waged sectors in 2014, as shown in Table 3.6. In additions, the disabled are dropped from the sample, since the nature may be different with the other workers.

Table 3.6: Sample Restrictions

Sample	**2000**		**2014/15**	
	Number of Jobs	% Lost	Number of Jobs	% Lost
Initial sample				
Age 14-110	25,825	-	36,381	-
Restrictions:				
Age 16-55	21,073	18.4	29,797	18.1
Employment: all employment status (tk01|tk02|tk03|tk04=1)	14,741	30.0	21,180	28.9
All employment status (tk01|tk02|tk03|tk04=1); including self-employment	12,580	14.7	16,672	21.3
Emp=1 and tk24a =4 and 5; 4: public workers; 5: private workers, excluding self-employed (only waged sector)	7,043	44.0	9,509	43.0
Education (non-missing) (tk01a==1 and tk01==1) and tk24a =4 and 5	6,718	4.6	9,423	0.9
Wage (non-missing)	6,634	1.3	9,139	3.0
Wage (non-zero)	6,537	1.5	9,008	1.4
Total sample	***6386***	***2.3***	***8119***	***9.9***
(control variables missing)				

Source: The author’s calculation.

3.3.3 Summary Statistics

Table 3.7 reports the summary statistics for the main variables and other control variables only for the sample in the present research. There were 6,386 observations in 2000 and 8,119 observations in 2014, respectively. For the dependent variables, in 2000, the minimum value of real hourly wages was IDR 0.0004287; it is worth noting that wage variable in the present study is self-reported, and the measurement error likely occurs, as Bauhoff (2014) asserted, self-report bias is the deviation between the self-reported and true values of the same measure. The bias is a type of measurement error that may be random or systematic and constant or variable. It can mislead descriptive statistics and causal inferences. While, the mean of real hourly wages was IDR 3,882.9 or GBP 0.29; this increased to around 2.9 times to IDR 16,046.8 or GBP 0.85 in 2014. The maximum hourly wages in 2014 was GBP 1,168.9 which was significantly different from the maximum hourly wage in 2000 (GBP 159.2). This increase in wage is in line with ILO's (2015) finding that the average net wages per month for regular employees (in real term, 2007=100) was around GBP 59.1 (IDR 1 million) in 2006, increased to GBP 61.5 (IDR 1.2 million) in 2014[20].

In addition, there are several methods to eliminate the outliers, such as: using the National Minimum Wage as the standard of bottom limit, for example: the minimum wage in 2014 was IDR 1,584,391, or the hourly wage was around IDR 9902.4 (GBP 0.54). The number of observations that have value less than 0.54 was 603 or around 2% of the sample. Alternative method of handling extreme scores (trimming) involves dropping the top and the bottom scores or some percentage of score, such as the top and bottom of 1% of scores (Warner, 2013).

[20] 1 GBP=IDR 16,909.6 in 2006; 1 GBP=IDR 19,503.2 in 2014.

Table 3.7: Summary Statistics, 2000 and 2014 Periods

	IFLS3								IFLS5							
						Percentiles								Percentiles		
	Obs	Mean	Std. Dev.	Min	Max	25%	50%	75%	Obs	Mean	Std. Dev.	Min	Max	25%	50%	75%
Dependent Variables																
Real Hourly wages (IDR)	6386	3882.9	26450.9	0.0	2000800.0	1020.8	1908.6	3841.5	8119	16046.8	254681.1	9.3	22700000.0	3591.3	6992.9	12968.7
Real Hourly wages (log)	6386	7.6	1.1	-7.8	14.5	6.9	7.6	8.3	8119	8.8	1.1	2.2	16.9	8.2	8.9	9.5
Independent Variables																
Main Variables																
Years of schooling (years)	6386	9.7	3.3	6.0	15.0	6.0	9.0	12.0	8119	11.5	3.4	6.0	22.0	9.0	12.0	15.0
Experience (years)	6386	15.8	10.3	0.0	43.0	7.0	14.0	23.0	8119	16.1	10.2	0.0	43.0	7.0	15.0	23.0
Other Control																
Job and Firm related																
Tenure (years)	6386	6.6	7.4	0.0	50.0	1.0	2.0	10.0	8119	6.5	7.3	0.0	40.0	1.3	4.0	9.2
Control Variables: Categorical Variables																
Education group	6386			1.0	4.0	1.0	2.0	3.0	8119			1.0	4.0	2.0	3.0	4.0

Personal Characteristics																
Sex	6386			0.0	1.0	0.0	0.0	1.0	8119			0.0	1.0	0.0	0.0	1.0
Marital Status	6386			1.0	3.0	1.0	2.0	2.0	8119			1.0	3.0	2.0	2.0	2.0
Religion	6386			1.0	5.0	1.0	1.0	1.0	8119			1.0	6.0	1.0	1.0	1.0
Ethnicity	6386			1.0	7.0	1.0	2.0	5.0	8119			1.0	7.0	1.0	2.0	7.0
Job and Firm related																
Full-time	6386			0.0	1.0	1.0	1.0	1.0	8119			0.0	1.0	1.0	1.0	1.0
Sector	6386			0.0	1.0	1.0	1.0	1.0	8119			2.0	3.0	3.0	3.0	3.0
Industry	6386			1.0	9.0	3.0	6.0	9.0	8119			1.0	10.0	3.0	6.0	9.0
Firm size	6386			1.0	4.0	1.0	2.0	3.0	8119			1.0	4.0	2.0	2.0	3.0
Regional																
Urban	6386			0.0	1.0	0.0	1.0	1.0	8119			0.0	1.0	0.0	1.0	1.0
Province	6386			12.0	73.0	31.0	33.0	35.0	8119			11.0	76.0	31.0	33.0	36.0

Source: The author’s calculation

For the dependent variables, Table 3.8 provides the distribution of education levels of the sample; most of the individuals in 2000 had primary school qualifications or below (36 per cent). In contrast, most of the sample had senior high school (39.5 per cent) and university qualifications (26.4 per cent) in 2014. There is a significant decline in workers with only primary school qualifications or below (-54 per cent); a possible reason for this change is the successful education reform and the compulsory education program. Before 2008, the compulsory education is up to 6 years (primary school only); after that, the program extends to 9 years of education (primary and junior high schools). The success of those programs is also confirmed by years of schooling data, as the mean value of years of schooling in 2000 was 9.7 years (around junior high school) which then increased to 11.5 years in 2014 (around senior high school). On the other hand, there was a significant rise in workers with university degrees, from 14.8 per cent to 26.4 per cent between the periods; this is in line with The British Council's (2012) finding that Indonesia's higher education participation grew by 53 per cent from 2000 to 2009 and is projected to continue growing.

Table 3.8: Sample Distribution Based on Education

Education Level	2000		2014	
	Freq.	**Per cent**	**Freq.**	**Per cent**
Group 1: Primary School or below	2,299	36.0	1,404	17.3
Group 2: Junior High School	1,162	18.2	1,360	16.8
Group 3: Senior High School	1,978	31.0	3,210	39.5
Group 4: University	947	14.8	2,145	26.4
Total	6,386	100.0	8,119	100.0

Source: The author's calculation.

The mean value of experience in 2000 is 15.8 years, which then increases slightly in 2014 to 16.1 years (Table 3.7). Nonetheless, the maximum, minimum and percentile of years of experience in both waves are similar, min. 0 year; max. 43 years; 25 per cent percentile: 7 years; 75 per cent percentile: 23 years, except for the 50 per cent percentile which was 14 years in 2000 but then increases to 15 years in 2014.

For the other control variables, firstly: sex, 66 per cent of the sample from 2000 was males and 34 per cent was females. Female increased to 39 per cent in 2014 (Table 3.9), possibly due to the increase of female participation in the economy, particularly in urban areas; as Schaner and Das (2016) find that younger women in urban areas have increased their labour force participation in recent years.

Table 3.9: Sample Distribution Based on Sex

Sex	2000		2014	
	Freq.	Per cent	Freq.	Per cent
Male	4,216	66.02	4,946	60.92
Female	2,170	33.98	3,173	39.08
Total	6,386	100	8,119	100

Source: The author's calculation.

Secondly, marital status: married and cohabitate status was 66.5 per cent and single was 29.5 per cent in 2000. Married and cohabitate individuals then increase by 13.3 per cent to 74.4 per cent in 2014 (Table 3.10). With regards to religion, 89.76 per cent in 2000 and 89.43 per cent in 2014 were Muslims (Table 3.11). This is in line with the fact that almost 90 per cent of the population in Indonesia is Muslims.

Table 3.10: Sample Distribution Based on Marital Status

Marital Status	2000		2014	
	Freq.	Per cent	Freq.	Per cent
Single	1,886	29.53	1,660	20.45
Married and Cohabitate	4,248	66.52	6,122	75.40
Other (widowed/divorced/separated)	252	3.95	337	4.15
Total	6,386	100	8,119	100

Source: The author's calculation.

Table 3.11: Sample Distribution Based on Religion

Religion	2000		2014	
	Freq.	Per cent	Freq.	Per cent
Religion: Islam	5,732	89.76	7,261	89.43
Religion2: Christian/Protestant	230	3.6	114	1.4
Religion3: Catholic	121	1.89	278	3.42
Religion4: Hindu	265	4.15	452	5.57
Religion5: Buddhist	38	0.6	14	0.17
Total	6,386	100	8,119	100

Source: The author's calculation.

In terms of ethnicity, Jawa turned out the largest in 2000 (46.8 per cent), Other was 22.1 per cent (omitted variable) and Sunda was 15.3 per cent. In 2014, the largest ethnicity remained the same; 45 per cent of Jawa, 23.5 per cent of Other, and 13.9 per cent of Sunda (Table 3.12).

Table 3.12: Sample Distribution Based on Ethnicity

Ethnicity Group	2000		2014	
	Freq.	Percent	Freq.	Percent
Jawa	2,989	46.81	3,664	45.13
Sunda	977	15.3	1089	13.41
Batak	169	2.65	311	3.83
Betawi	443	6.94	41	0.5
Minang	346	5.42	680	8.38
Tiong Hoa	46	0.72	43	0.53
Other	1,416	22.17	2,291	28.22
Total	6,386	100	8,119	100

Source: The author's calculation.

Turning to work and firm size, more than 85 per cent of workers are full-time workers. Part-time workers made up 14.5 per cent of the sample in 2000 and 16.90 per cent in 2014 (Table 3.13). Although the majority of sample is full-time workers, there is a rise in the number of part time workers (people who work less than 30 hours per week) in the sample. In additions, part time is defined as workers in the waged sector who work less than 30 hours per week that follows The ILO Part-Time Work Convention, 1994 (No. 175), part time workers could work both in public and private sector.

The mean value for tenure in 2000 was 6.6 year which then decreased slightly in 2014 to 6.5 years[21] (Table 3.8). If the sample is separated by gender, the mean value became 6.3 years for males' tenure and 5.8 years for females' tenure. This average was the same for both 2000 and 2014. For sector dummy, public sector was only around 16 per cent of the total sector in both waves. Also, the sample was dominated by private sectors, *i.e.* more than 80 per cent (Table 3.14). A possible reason for the decrease in tenure is that many employees work on short-term contracts. Allen (2016) confirms that the main form of labour market flexibility in the formal economy in Indonesia is through short-term contracting or outsourcing arrangements. However, according to the Manpower Law 13/2003, short-term contracts cannot be provided for work that is permanent in nature

[21] There is a difference between the maximum number of tenures in 2000 (50 years) and the maximum experience (43 years), this is due to the definition of variables. The present study uses potential experience and actual tenure based on IFLS question in the estimation. As a result, there are negative values of tenure. According to ACAPS (2016), there are common mistakes in needs assessments in the data, one of the mistakes is values are outside of the acceptable range of values for that question, including negative values in fields that can only have positive values. In addition, the number of respondents with tenure of more than 40 years is 21 observations or only 0.14 per cent of total sample; thus, those negative values are set to zero.

and short-term work agreements can only be made for an initial period of two years with an option to extend for an additional 12 months.

Table 3.13: Sample Distribution Based on Working Hours

Employment Status	2000			2014		
	Freq.	**Per cent**	**Cum.**	**Freq.**	**Per cent**	**Cum.**
Part time	929	14.55	14.55	1,372	16.90	16.90
Full time	5,457	85.45	100	6,747	83.10	100
Total	6,386	100.00		8,119	100.00	

Source: The author's calculation.

Table 3.14: Sample Distribution Based on Sector

Sector	2000		2014	
	Freq.	**Per cent**	**Freq.**	**Per cent**
Public	1,060	16.6	1,391	17.13
Private	5,326	83.4	6,728	82.87
Total	6,386	100	8,119	100

Source: The author's calculation.

In terms of industry, most of the respondents in 2000 was from social services (34.6 per cent), manufacturing (21.5 per cent), and agriculture (15.2 per cent). In 2014, the top three sectors were social services (36.7 per cent), manufacturing (21.9 per cent), and wholesale, retail, restaurants and hotels (17.5 per cent), as shown in Table 3.15. This change confirms that structural change indeed took place in Indonesia, shifting from the agricultural sector to the service sector.

Table 3.15: Sample Distribution Based on Industry

Industry	2000		2014	
	Freq.	**Per cent**	**Freq.**	**Per cent**
Agriculture	968	15.16	553	6.81
Mining and quarrying	56	0.88	141	1.74
Manufacturing	1,374	21.52	1,772	21.83
Electricity, gas and water	39	0.61	76	0.94
Construction	534	8.36	342	4.21
Wholesale, retail, restaurants and hotels	764	11.96	1,408	17.34
Transportation, storage, and communications	340	5.32	224	2.76
Finance, insurance, real estate and business services	103	1.61	522	6.43
Social services	2,208	34.58	2,975	36.64
Other		0.00	106	1.31
Total	6,386	100.00	8,119	100.00

Source: The author's calculation.

Furthermore, most of the industries in both periods were micro (1-4 workers) and small firms (5-19 workers). There is an increasing number of medium and large firms; Particularly for large firms, there was a significant rise from less than 10 per cent in 2000 to more than 20 per cent in 2014 (Table 3.16).

Table 3.16: Sample Distribution Based on Firm-Size (Number of Workers)

Number of workers	2000		2014	
	Freq.	**Percent**	**Freq.**	**Percent**
1-4 workers	2,215	34.69	1,848	22.76
5-19 workers	2,218	34.73	2,459	30.29
20-99 workers	1,351	21.16	2,069	25.48
≥100	602	9.43	1,743	21.47
Total	6,386	100	8,119	100

Source: The author's calculation.

The last category is regional/residence dummy; 63 per cent of the sample in 2000 lived in urban areas, this then increased to 71.65 per cent in 2014. This implies that most of the respondents lived in urban areas (Table 3.17). Meanwhile for provinces, most respondents in 2000 lived in Jawa Barat (19.9 per cent), Jawa Timur (14.3 per cent), and DKI Jakarta – the capital region (13.2 per cent); a similar trend also persists for the 2014 period. This also confirms that most Indonesian population lived in Jawa island - the most populous and developed island in Indonesia (Table 3.18).

Table 3.17: Sample Distribution Based on Rural/Urban Area

Residency	2000		2014	
	Freq.	Per cent	Freq.	Per cent
Rural	2,364	37.02	2,302	28.35
Urban	4,022	62.98	5,817	71.65
Total	6,386	100	8,119	100

Source: The author's calculation.

Table 3.18: Sample Distribution Based on Province

Province	2000		2014	
	Freq.	Per cent	Freq.	Per cent
Aceh			1	0.01
Sumatera Utara	337	5.28	485	5.97
Sumatera Barat	291	4.56	389	4.79
Riau	40	0.63	60	0.74
Jambi		0.00	12	0.15
Sumatera Selatan	211	3.30	311	3.83
Lampung	162	2.54	229	2.82
Kepulauan Bangka Belitung		0.00	80	0.99
Kepulauan Riau		0.00	27	0.33
DKI Jakarta	843	13.20	702	8.65
Jawa Barat	1,271	19.90	1,178	14.51
Jawa Tengah	822	12.87	1030	12.69
D I Yogyakarta	414	6.48	469	5.78
Jawa Timur	915	14.33	1005	12.38
Banten		0.00	377	4.64
Bali	296	4.64	490	6.04
Kalimantan Barat		0.00	521	6.42
Kalimantan Tengah	306	4.79	4	0.05
Kalimantan Selatan	7	0.11	15	0.18
Kalimantan Timur	242	3.79	341	4.20
Sulawesi Selatan	4	0.06	52	0.64
Sulawesi Tenggara	225	3.52	329	4.05
Papua Barat			12	0.15
Total	6,386	100.00	8,119	100.00

Source: The author's calculation.
Note: DKI Jakarta is the capital city.

3.4 Estimation Result

3.4.1 Basic Mincer Equation

Based on the estimation of the basic Mincer equation (Table 3.19), the marginal return (hourly wage) to complete each additional level of education (relative to primary school or below) increases alongside the level of education in both 2000 and 2014 periods. Furthermore, the present study provides other Mincer equation specifications, a model with basic and personal characteristics, and a model with basic, personal characteristics, work related and firm size. The estimation result is provided in Appendix VI. This part will elaborate the result for main variables, i.e., education level and experience. While, the control variables analysis is provided in Section 3.4.6.

Table 3.19: Estimation Result: Basic Mincer Wage Equation, All Individuals, 2000 and 2014 Periods

	Basic Regression					
	2000			2014		
Log of real hourly wage	**Coef.**	**SE**	**P>t**	**Coef.**	**SE**	**P>t**
Junior High School	0.427	0.035	***	0.326	0.040	***
Senior High School	0.899	0.030	***	0.755	0.035	***
University	1.597	0.037	***	1.268	0.038	***
Experience	0.059	0.004	***	0.044	0.004	***
Experience squared	-0.001	0.000	***	-0.001	0.000	***
Constants	6.369	0.040	***	7.649	0.044	***
Observation		6386			8119	
the R-squared statistic		0.260			0.153	

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

In 2000, the highest return to education occurred in the university level; 159.7 per cent higher than primary school or below, and the lowest was the return to junior high school (42.7 per cent) relative to the base group. Experience coefficient was positive and significant (0.059), implying that an additional year of experience will increase hourly wage by 5.9 per cent; thus, hourly wages grow as a (concave) function of experience. Also, coefficient of experience squared was negative and significant (-0.001); this confirms that there is a diminishing return on experience, as stated in Section 3.2.2. All of the variables are significantly different from zero or statistically significant at 1 per cent. The R-squared statistic for this model is 0.260 per cent which implies that 26 per cent of the variation in log of hourly wage is explained by the variation in education level, experience and experience squared (the standard Mincer equation). These results come from 6,386 observations.

Similar pattern of return to education from the year 2000 continues to 2014; the highest return was at the university level (126.8 per cent higher than the base group) and the lowest return was at the junior high school (32.6 per cent higher than primary school or below). The experience coefficient was 0.044 (positive and significant) and the experience square coefficient was -0.001 (negative and significant). The R-squared statistic is 0.153 or only 15.3 per cent with the number of observations being 8,119. These findings from both years confirm the human capital and signalling theories and previous empirical studies (Montenegro and Patrinos, 2013; Purnastuti *et al.*, 2013; and Dong, 2016), *i.e.* the returns to schooling in Indonesia tend to increase as the level of education increases.

Furthermore, adding control variables does not change the pattern of the main variables; the return to education still increases in line with the education level. Nonetheless, each coefficient value decreases slightly compared to the basic model; for instance, the return to university level was 119.3 per cent in 2000 and 95.8 per cent in 2014 (Table 3.20). The coefficients are the measurement of how much each variable explains the response variable; adding more correlated variables causes the effect to overlap in explaining the response, thus adding more variables will affect each coefficient. Besides, including more coefficients also boosts the R-squared statistic; as explanatory variables are added to the model, each explanatory variable will explain some of the variance in the dependent variable (log of wages) as simply due to chance. One could continue to add predictors to the model, which would continue to improve the ability of the predictors to explain the dependent variable. In fact, R Square will always increase when adding more predictors. However, some of this increase in R Square is simply probably because of chance variation in that particular sample; it would not necessarily mean that the model gets better. Correspondingly, Ryznar and Vidican (2006) argue that the adjusted R Square attempts to yield a more honest value to estimate the R Squared for the population.

Compared to previous empirical studies, the result from the present study is more similar to that of Gropello and Sakellariou (2010) rather than other studies in Indonesia (as explained in Section 3.2.5), although the present study has a different set of data sources, time period and model specifications. The difference between the return to high school and junior high school was around 25-35 per cent, and the difference between return to university and senior high school was around 50 per cent during 2000 and 2014 periods.

Table 3.20: Estimations Result: Mincer Wage Equation with Control Variables, All Individuals, 2000 and 2014 Periods

	2000			2014		
	All			All		
Log of real hourly wage	**Coef.**	**SE**	**P>t**	**Coef.**	**SE**	**P>t**
Junior High School	0.285	0.034	***	0.221	0.037	***
Senior High School	0.637	0.033	***	0.480	0.035	***
University	1.193	0.044	***	0.958	0.043	***
Experience	0.032	0.005	***	0.022	0.004	***
Experience squared	-0.001	0.000	***	0.000	0.000	***
Sex (1=female)	-0.345	0.025	***	-0.279	0.022	***
Married and cohabitate	0.099	0.033	***	0.048	0.033	
Other (Separated, divorced and widowed)	-0.048	0.064		0.101	0.059	*
Religion1: Islam	-0.011	0.106		0.069	0.091	
Religion2: Christian/Protestant	0.079	0.121		0.232	0.123	*
Religion3: Catholic	0.048	0.127		0.207	0.106	*
Religion5: Buddhist	-0.195	0.198		0.156	0.303	
Ethnicity1: Jawa	-0.039	0.047		0.117	0.033	***
Ethnicity2: Sunda	-0.062	0.057		-0.123	0.373	
Ethnicity3: Batak	0.073	0.094		0.011	0.045	
Ethnicity4: Betawi	0.014	0.066		0.287	0.167	*
Ethnicity5: Minang	0.038	0.093		-0.045	0.912	
Ethnicity6: Tiong Hoa	0.121	0.157		0.175	0.323	
Status: full-time (30 hours a week or more)	-0.636	0.032	***	-0.461	0.029	***
Tenure	0.021	0.005	***	0.019	0.004	***
Tenure squared	0.000	0.000		0.000	0.000	
Sector: private	-0.291	0.038	***	-0.286	0.034	***
Industry2: mining and quarrying	0.373	0.120	***	0.383	0.087	***
Industry3: manufacturing	0.139	0.040	***	0.053	0.045	
Industry4: electricity, gas and water	0.139	0.143		0.155	0.112	
Industry5: construction	0.339	0.048	***	0.307	0.062	***
Industry6: wholesale, retail, restaurants and hotels	0.097	0.046	**	-0.018	0.046	
Industry7: transportation, storage, and communications	0.187	0.057	***	0.115	0.072	
Industry8: Finance, insurance, real estate and business services	0.432	0.093	***	0.235	0.057	***
Industry9: Social services	0.119	0.038	***	-0.130	0.044	***
Firm size2: 5-19 people	0.183	0.027	***	0.067	0.029	**
Firm size3: 20-99 people	0.243	0.032	***	0.246	0.031	***
Firm size4: >= 100 people	0.353	0.043	***	0.525	0.035	***
Province1: Aceh		(omitted)		1.285	0.918	
Province2: Sumatera Utara	0.372	0.085	***	0.425	0.061	***
Province3: Sumatera Barat	0.273	0.110	**	0.518	0.062	***
Province4: Riau	1.075	0.156	***	0.808	0.126	***
Province5: Jambi		(omitted)		0.580	0.268	**
Province6: Sumatera Selatan	0.207	0.080	***	0.368	0.068	***
Province7: Lampung	0.153	0.092	*	0.170	0.076	**

Province 8: Kepulauan Bangka Belitung		(omitted)		0.599	0.110	***
Province 9: Kepulauan Riau		(omitted)		0.966	0.181	***
Province10: DKI Jakarta	0.538	0.076	***	0.694	0.057	***
Province11: Jawa Barat	0.412	0.074	***	0.434	0.060	***
Province12: Jawa Tengah	0.156	0.074	**	0.108	0.059	*
Province13: D I Yogyakarta	0.018	0.081		0.069	0.067	
Province14: Jawa Timur	0.176	0.070	**	0.139	0.057	**
Province15: Banten		(omitted)		0.632	0.067	***
Province16: Bali	0.147	0.111		0.609	0.094	***
Province17: Kalimantan Barat		(omitted)		1.329	0.460	***
Province18: Kalimantan Tengah	0.780	0.332	**	1.309	0.241	***
Province19: Kalimantan Selatan	0.336	0.075	***	0.523	0.065	***
Province20: Kalimantan Timur	0.912	0.437	**	0.943	0.134	***
Province21: Sulawesi Selatan	0.010	0.077		0.323	0.065	***
Province22: Sulawesi Tenggara		(omitted)		0.217	0.266	
Urban	0.002	0.026		0.147	0.024	***
Constants	7.040	0.132	***	8.471	0.154	***
Observation	6386			8119		
The R-squared statistic	0.372			0.301		

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

3.4.2 Mincer Wage Equation by Gender

In this part, the analysis of basic Mincer equation is separated by gender. The pattern of coefficient in the estimations does not change and the return to education (in the standard regression) increases in line with the level of education for both males and females. Nevertheless, in terms of the coefficient of education level, the result shows that the females' coefficient is higher than that of the males', implying that the return to education for females is higher than that for males, though the result could be overestimated since females have shorter working hours and experience, as explained in Chapter 2.

It is worth noting as well that the present study uses potential experience as experience variable. Potential experience is likely to be an overestimate of true experience, particularly for women, since women may have shorter work experience and working hours as discussed in Chapter 2. This is not picked up by the Mincer equation or by Heckman selection; as a result, an upward bias could occur in Mincer wage equation, and cautious interpretation of return to schooling is required particularly related to the gender gap. The finding is in line with Abbasa and Foreman-Peck (2008), who assert that high measured returns to females' education is caused by a combination of much lower workforce participation and fewer average years in the labour force than males. Thus, the actual returns are much lower. Presumably, higher returns are necessary to compensate for the shorter period and lower probability of wages.

In 2000, the highest return to education for males was at the university level (142.5 per cent) and the lowest return was at junior high school (around one-fourth the return to university). The effect of experience on wages was 5.8 per cent and was diminishing. On the other hand, the return of females was significantly higher *i.e*. 40.7 per cent for junior high school relative to the primary school or below and 186.7 per cent for university graduates. The effect of experience on female's wages was slightly lower than on males; only 5.4 per cent, with the R-squared statistic of females slightly higher than that of males. This implies that the combination of education level and experience have more effects on females than on males. The estimations also show similar pattern in 2014, however with lower value of coefficient for both males and females (Table 3.21); this trend is also similar to estimations for all individuals (the sample).

Table 3.21: Estimation Result: Basic Mincer Equation by Gender, 2000 and 2014/16 Periods

	Gender											
	2000						2014					
	Male			Female			Male			Female		
Log of real hourly wage	Coef.	SE	P>t	Coef.	SE	P>t	Coef.	SE	P>t	Coef.	SE	P>t
Junior High School	0.355	0.040	***	0.407	0.066	***	0.268	0.048	***	0.381	0.068	***
Senior High School	0.734	0.036	***	1.127	0.055	***	0.654	0.042	***	0.823	0.062	***
University	1.425	0.046	***	1.867	0.061	***	1.213	0.047	***	1.386	0.062	***
Experience	0.058	0.005	***	0.054	0.007	***	0.043	0.005	***	0.037	0.006	***
Experience squared	-0.001	0.000	***	-0.001	0.000	***	-0.001	0.000	***	0.000	0.000	***
Constants	6.571	0.049	***	6.113	0.068	***	7.845	0.055	***	7.458	0.073	***
Observation		4216			2170			4946			3173	
The R-squared statistic		0.22			0.34			0.15			0.17	

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent;
*** significance level at 1 per cent.

Adding control variables, as shown in Table 3.22, also does not change the pattern of coefficient for both genders (the other model specifications are provided in Appendix VI) although the value is slightly lower than basic regression result. The highest return to education is still at the university level and the lowest is at junior high school, relative to primary school or below for both genders and both periods. Experience squared for females is insignificant in both periods; implying that experience squared does not have a diminishing return for females.

Additionally, the present study run the other estimations using the interaction variables between gender and years of schooling, as provided in Appendix VI.5, the coefficient of the interaction term is positive and significant affect wages. Thus, there is a combination effects of gender – years of schooling and the wages.

In short, the present study finds that the return to education is relatively higher for females than for males. This is in line with previous empirical studies such as: Psacharopoulos (1994), Harmon *et al.* (2000), and Dumauli (2015), as explained in Section 3.2.3. Dougherty (2005) also furthers that there are some possible explanations for gender differences in the return to education, one of which is female advantages, for example: garment companies prefer to recruit females. The higher return for females is affected by both of supply and demand side. From the demand side, the changes are in line with economic transformation and improvements that increases the opportunity of female workers to participate in the labour market. From the supply side, there is an increase in education participation of females. Finally, education increases female's skills and productivity, as it does with male, and in addition it appears to reduce the gap in male and female earnings attributable to factors such as discrimination, tastes, and circumstances (Dougherty, 2005).

Table 3.22: Estimation Result: Mincer Wage Equation with Control Variables, by Gender

	Female						Male					
	2000			2014			2000			2014		
Log of real hourly wage	**Coef.**	**SE**	**P>t**	**Coef.**	**SE.**	**P>t**	**Coef.**	**SE.**	**P>t**	**Coef.**	**SE**	**P>t**
Junior High School	0.251	0.063	***	0.308	0.063	***	0.281	0.040	***	0.176	0.045	***
Senior High School	0.770	0.060	***	0.586	0.061	***	0.572	0.040	***	0.423	0.042	***
University	1.311	0.077	***	1.019	0.074	***	1.101	0.055	***	0.899	0.053	***
Experience	0.017	0.008	**	0.022	0.007	***	0.044	0.006	***	0.024	0.005	***
Experience squared	0.000	0.000		0.000	0.000	**	-0.001	0.000	***	0.000	0.000	***
Married and cohabitate	0.085	0.055		-0.089	0.055		0.121	0.042	***	0.135	0.040	***
Other (Separated, divorced and widowed)	0.022	0.083		0.001	0.083		0.057	0.120		0.164	0.091	*
Religion1: Islam	0.118	0.173		0.124	0.147		-0.077	0.134		0.028	0.114	
Religion2: Christian/Protestant	0.318	0.193	*	0.354	0.196	*	-0.074	0.153		0.141	0.156	
Religion3: Catholic	0.224	0.213		0.372	0.171	**	-0.073	0.157		0.071	0.134	
Religion5: Buddhist	-0.102	0.345		0.079	0.547		-0.216	0.240		0.163	0.359	
Ethnicity1: Jawa	-0.063	0.085		0.140	0.057	**	-0.032	0.056		0.117	0.040	***
Ethnicity2: Sunda	0.014	0.102		(omitted)			-0.125	0.069	*	-0.134	0.360	
Ethnicity3: Batak	-0.078	0.164		0.015	0.077		0.136	0.114		0.017	0.054	
Ethnicity4: Betawi	0.142	0.114		0.290	0.303		-0.090	0.081		0.275	0.198	
Ethnicity5: Minang	0.046	0.160		(omitted)			-0.007	0.113		0.019	0.880	
Ethnicity6: Tiong Hoa	0.033	0.264		-0.433	0.947		0.185	0.193		0.395	0.333	
Status: fulltime (30 hours a week or more)	-0.584	0.049	***	-0.361	0.043	***	-0.722	0.042	***	-0.622	0.040	***
Tenure	0.029	0.008	***	0.025	0.007	***	0.013	0.006	**	0.014	0.005	***
Tenure squared	0.000	0.000		0.000	0.000		0.000	0.000		0.000	0.000	
Sector: private	-0.436	0.068	***	-0.338	0.056	***	-0.218	0.045	***	-0.254	0.044	***
Industry2: mining and quarrying	1.213	0.500	**	0.501	0.367		0.324	0.124	***	0.438	0.089	***
Industry3: manufacturing	0.168	0.071	**	-0.195	0.085	**	0.159	0.048	***	0.173	0.053	***
Industry4: electricity, gas and water	0.229	0.358		-0.331	0.485		0.144	0.154		0.272	0.114	**
Industry5: construction	0.063	0.166		0.447	0.226	**	0.316	0.052	***	0.319	0.066	***
Industry6: wholesale, retail, restaurants and hotels	0.153	0.078	**	-0.236	0.086	***	0.083	0.057		0.097	0.055	*
Industry7: transportation, storage, and communications	-0.174	0.221		0.155	0.231		0.181	0.060	***	0.169	0.075	**
Industry8: Finance, insurance, real estate and business services	0.544	0.157	***	0.138	0.113		0.361	0.115	***	0.291	0.064	***
Industry9: Social services	0.134	0.070	*	-0.287	0.083	***	0.164	0.046	***	-0.045	0.052	

Firm size2: 5-19 people	0.333	0.049	***	0.021	0.048		0.105	0.033	***	0.080	0.036	**
Firm size3: 20-99 people	0.419	0.056	***	0.285	0.053	***	0.142	0.039	***	0.206	0.038	***
Firm size4: >= 100 people	0.590	0.069	***	0.636	0.060	***	0.232	0.054	***	0.454	0.042	***
Province1: Aceh		(omitted)			(omitted)			(omitted)		1.268	0.886	
Province2: Sumatera Utara	0.618	0.148	***	0.436	0.104	***	0.290	0.103	***	0.430	0.075	***
Province3: Sumatera Barat	0.218	0.192		0.739	0.100	***	0.329	0.133	**	0.376	0.078	***
Province4: Riau	1.288	0.310	***	0.998	0.272	***	0.967	0.179	***	0.772	0.140	***
Province5: Jambi		(omitted)		0.724	0.553			(omitted)		0.428	0.299	
Province6: Sumatera Selatan	0.241	0.147		0.518	0.123	***	0.200	0.094	**	0.299	0.080	***
Province7: Lampung	0.062	0.176	*	0.101	0.136		0.169	0.107		0.184	0.090	**
Province8: Kepulauan Bangka Belitung		(omitted)		0.733	0.184	***		(omitted)		0.521	0.135	***
Province9: Kepulauan Riau		(omitted)		1.262	0.395	***		(omitted)		0.907	0.200	***
Province10: DKI Jakarta	0.737	0.133	***	0.803	0.099	***	0.460	0.092	***	0.633	0.069	***
Province11: Jawa Barat	0.510	0.131	***	0.531	0.102	***	0.389	0.088	***	0.383	0.073	***
Province12: Jawa Tengah	0.203	0.131		0.116	0.099		0.142	0.089		0.110	0.073	
Province13: D I Yogyakarta	0.096	0.140		0.177	0.112		0.016	0.098		0.011	0.083	
Province14: Jawa Timur	0.326	0.126	***	0.177	0.096	*	0.120	0.083		0.133	0.069	*
Province15: Banten		(omitted)		0.890	0.112	***		(omitted)		0.458	0.083	***
Province16: Bali	0.533	0.185	***	0.851	0.155	***	-0.020	0.138		0.468	0.118	***
Province17: Kalimantan Barat		(omitted)		1.635	0.678	**		(omitted)		1.054	0.625	*
Province18: Kalimantan Tengah	1.431	0.616	**	0.909	0.431	**	0.628	0.388		1.495	0.285	***
Province19: Kalimantan Selatan	0.637	0.142	***	0.442	0.113	***	0.228	0.088	***	0.552	0.078	***
Province20: Kalimantan Timur	0.481	0.876		0.815	0.237	***	1.041	0.499	**	1.021	0.161	***
Province21: Sulawesi Selatan	0.204	0.135		0.174	0.106		-0.085	0.093		0.451	0.081	***
Province22: Sulawesi Tenggara		(omitted)		-0.275	0.551			(omitted)		0.335	0.297	
Urban	-0.037	0.047		0.187	0.041	***	0.027	0.031		0.126	0.030	***
Constants	6.438	0.224	***	8.250	0.250	***	7.156	0.162	***	8.545	0.195	***
Observation	2170			3173			4216			4946		
The R-squared statistic	0.471			0.332			0.314			0.285		

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

3.4.3 Mincer Wage Equation by Sectors

This part elaborates on the analysis based on sectors (public and private sectors). As explained in Section 2.1.3, wage determination in public and private sector is different. Wages in private sector in Indonesia are largely determined by the market. On the other hand, the determination of public service's wage is more complicated, involving seniority, position, rank, and political approach. Thus, the hypothesis is the return to education in public sector is relatively higher than in private sector, as Psacharopolous (1979) and Chevalier et al. (2004) findings.

The result of basic Mincers equation for public and private sectors in 2000 and 2014 shows a consistency; the return to education increases in line with the level of education and the highest return is at the university level, as shown in Table 3.23. Also, experience has a positive and significant effect and experience squared has a positive and diminishing effect. The R-squared statistic is higher for public sector than for private sector, however, the R-squared statistic for both sectors tend to decrease in 2014, possibly because more variables which are accounted for in the present study affect wage equation in certain ways following recent developments.

Table 3.23: Estimation Result: Basic Mincer Wage Equation by Sector

Log of real hourly wage	Sector											
	2000						2014					
	Public			Private			Public			Private		
	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>t
Junior High School	0.413	0.126	***	0.375	0.037	***	0.180	0.250		0.254	0.040	***
Senior High School	0.961	0.100	***	0.789	0.034	***	0.852	0.205	***	0.629	0.036	***
University	1.532	0.103	***	1.474	0.051	***	1.484	0.202	***	1.027	0.043	***
Experience	0.052	0.010	***	0.054	0.005	***	0.084	0.011	***	0.034	0.004	***
Experience squared	0.000	0.000	***	-0.001	0.000	***	-0.001	0.000	***	-0.001	0.000	***
Constants	6.450	0.130	***	6.470	0.045	***	7.015	0.212	***	7.888	0.046	***
Observation		1060			5326			1391			6728	
The R-squared statistic		0.264			0.166			0.229			0.096	

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

When comparing both sectors, the result shows that most of the coefficients' value in the public sector are higher than in the private sector, except for the return to junior high school in 2014 which was only 18 per cent in the public sector (but insignificant in contrast to return to primary school and below). This is in line with Filmer and Lindauer

(2001), Sharma (2013), and Taniguchi and Tuwo (2014) who assert that public sector in Indonesia receives substantially higher wages than other sectors. The distribution of earnings across the public sector suggests that government wage-setting institutions are very different from those of the private economy. A possible argument for this is that the public sector prefers workers with high education level, with at least senior high school qualifications. For unskilled jobs such as cleaning, the public sector prefers outsourcing or using private firms' services, rather than hiring directly[22]. That being said, our finding is in contrast to that of Byron and Takahashi (1989)[23].

Some possible reasons for higher return in the public sector are: (1) civil servants' salary has increased significantly in real terms and at a much faster rate than those in the private sector since Indonesia's economy recovered from the 1997 financial crisis (The World Bank, 2000); (2) the majority of the sample is urban settlers with more than 60 per cent for both periods (Table 3.17), who tend to earn higher wages than those who settle in rural employment; (3) central government employees are covered by a unified salary structure that do not differentiate on the basis of rural or urban location (Filmer and Lindauer, 2001); and (4) private sector is more efficient than public sector in terms of productive and allocative efficiency (Rao, 2015).

Furthermore, Psacharopolous (1979) who studies a similar distinction between the public and private sectors argues that wages could exceed productivity in the public sector but not in the competitive private sector. The lack of competition in the public sector allows higher returns to education. Chevalier *et al.* (2004) also add that higher returns in the public sector support some signalling values of education.

Turning to the effect of experience on wages, the private sector had a slightly higher experience effect in 2000, *i.e.* 5.2 per cent for public sector and 5.4 per cent for private sector. In contrast, experience had higher effect on wages in the public sector (8.4 per cent) than on private sector (3.4 per cent) during 2014. Both sectors had a diminishing

[22] This is in line with The Minister of Manpower Regulation Number 19 of 2012 on the Terms to Outsource Work to Another Company' (Regulation 19). The regulation states that outsourcing activities may only be implemented in the following lines of work, *i.e.* janitorial work (or cleaning services), catering services, security guard services, supporting services in the mining and oil sectors; and transportation services.

[23] Byron and Takahashi (1989) study the return to education in the government and private sectors of urban Java, using the 1981 SUSENAS data. The return to the private sector is more generous than the return to the public sector. The general returns to education were of the order of 15-17 per cent additional income for each additional year of schooling. In comparison, the return to education in Indonesia was relatively higher than other countries in 1981.

return of experience (the coefficient of experience squared is negative and significant at 1 per cent). This is possibly related to the significant increase of education and experience levels in the public sector. In contrast, there is a slight decrease of experience in the private sector.

Furthermore, adding control variables does not change the pattern of return to education; the result shows that the coefficients are lower compared to the basic model (Table 3.24) - the other model specifications are provided in Appendix VI. Adding more correlated variables causes an overlap effect in explaining the response, thus the addition will affect each coefficient, as explained in the basic regression part. Those values, however, still imply that the returns to education increases in line with the education level, as in all individuals and estimation by gender.

In terms of experience variables and its squared in public sector, those variables are insignificant in the year 2000, suggesting that experience could not have any effect on return to education, as well as do not have a diminishing effect on wages. One of possible reason is insignificant different of wages between workers with 1-year experience and workers with 0-year experience, even though they have the same education qualification. As asserted in the Government Regulation 6/1997 (effective since 2000), for example: wages increase that allowed by the government, for rank level 3 (entry level for an undergraduate degree) with 0 to 1-year experience is IDR 241,800 per month (around GBP 18.4); for those with 2-year experience the wages increase is IDR 253,900 per month (around GBP 19.3). And this is consistent with the estimation result without experience variables, as indicated in Appendix VIII.

In contrast to public sector, experience and its squared significantly affect the wages in private sector in 2000. And both of them turn insignificant in 2014. One of possible reasons is the Minimum Wage Policy (MWP), as stated in the Ministry of Labour and Transmigration's regulation No. 7/2013. This law becomes compulsory for all private sectors, the adoption of the MWP was dominated by various forms of politicization that made irrational wage increases and caused uncertainty. Rather than considers workers' experience, the employers attempt to fulfil the MWP.

Table 3.24: Estimation Result: Mincer Wage Equation with Control Variables, by Sectors

	Private						Public					
	2000			2014			2000			2014		
Log of real hourly wage	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>t
Junior High School	0.259	0.036	***	0.179	0.037	***	0.335	0.129	***	0.168	0.240	
Senior High School	0.577	0.036	***	0.414	0.035	***	0.825	0.107	***	0.672	0.202	***
University	1.177	0.054	***	0.826	0.045	***	1.290	0.113	***	1.326	0.206	***
Experience	0.032	0.005	***	0.007	0.004		0.006	0.013		0.062	0.013	***
Experience squared	-0.001	0.000	***	0.000	0.000		0.000	0.000		-0.001	0.000	**
Sex (1=female)	-0.409	0.028	***	-0.304	0.024	***	-0.054	0.053		-0.143	0.057	**
Married and cohabitate	0.097	0.036	***	0.069	0.034	**	0.236	0.096	**	0.294	0.103	***
Other (Separated, divorced and widowed)	-0.058	0.070		0.101	0.062		0.182	0.171		0.536	0.178	***
Religion1: Islam	-0.050	0.127		0.015	0.098		0.047	0.178		0.329	0.226	
Religion2: Christian/Protestant	0.084	0.145		0.283	0.133	**	-0.018	0.203		0.224	0.296	
Religion3: Catholic	0.092	0.151		0.222	0.115	*	-0.233	0.220		0.196	0.258	
Religion5: Buddhist	-0.215	0.213		0.068	0.304		-0.017	0.807		1.583	1.002	
Ethnicity1: Jawa	-0.045	0.053		0.098	0.035	***	0.023	0.095		0.154	0.089	*
Ethnicity2: Sunda	-0.054	0.065		-0.172	0.362		-0.141	0.122			(omitted)	
Ethnicity3: Batak	0.095	0.106		-0.008	0.046		0.080	0.216		-0.017	0.140	
Ethnicity4: Betawi	0.030	0.073		0.244	0.163		-0.303	0.179	*		(omitted)	
Ethnicity5: Minang	-0.007	0.106		0.092	0.885		0.112	0.183			(omitted)	
Ethnicity6: Tiong Hoa	0.102	0.161		0.277	0.362			(omitted)		-0.381	0.695	
Status: fulltime (30 hours a week or more)	-0.687	0.036	***	-0.514	0.033	***	-0.464	0.068	***	-0.341	0.064	***
Tenure	0.020	0.005	***	0.037	0.005	***	0.041	0.011	***	-0.002	0.012	
Tenure squared	0.000	0.000	*	-0.001	0.000	***	-0.001	0.000	**	0.001	0.000	
Industry2: mining and quarrying	0.339	0.125	***	0.373	0.090	***	0.801	0.552		0.440	0.282	
Industry3: manufacturing	0.127	0.042	***	0.072	0.046		0.083	0.150		-0.375	0.279	

Industry4: electricity, gas and water	0.179	0.183		0.247	0.118	**	0.016	0.224		-0.416	0.324	
Industry5: construction	0.342	0.051	***	0.331	0.062	***	0.155	0.188		-0.030	0.511	
Industry6: wholesale, retail, restaurants and hotels	0.099	0.048	**	-0.006	0.047		-0.150	0.207		-0.273	0.405	
Industry7: transportation, storage, and communications	0.188	0.061	***	0.119	0.072		0.212	0.156		0.643	0.374	*
Industry8: Finance, insurance, real estate and business services	0.476	0.104	***	0.260	0.058	***	0.307	0.202		0.158	0.206	
Industry9: Social services	0.087	0.042	**	-0.123	0.046	***	0.157	0.102		-0.170	0.143	
Firm size2: 5-19 people	0.172	0.029	***	0.059	0.030	**	0.169	0.072	**	0.128	0.104	
Firm size3: 20-99 people	0.255	0.035	***	0.231	0.033	***	0.194	0.077	**	0.327	0.104	***
Firm size4: >= 100 people	0.418	0.047	***	0.531	0.036	***	0.133	0.107		0.596	0.116	***
Province1: Aceh	(omitted)			1.126	0.892		(omitted)			(omitted)		
Province2: Sumatera Utara	0.480	0.096	***	0.373	0.070	***	0.050	0.204		0.609	0.133	***
Province3: Sumatera Barat	0.357	0.127	***	0.417	0.073	***	0.077	0.211		0.664	0.118	***
Province4: Riau	1.233	0.168	***	0.662	0.135	***	0.045	0.471		1.297	0.333	***
Province5: Jambi	(omitted)			0.504	0.261	*	(omitted)			(omitted)		
Province6: Sumatera Selatan	0.272	0.091	***	0.321	0.076	***	0.054	0.159		0.342	0.152	**
Province7: Lampung	0.208	0.103	**	0.097	0.082		0.107	0.213		0.243	0.195	
Province8: Kepulauan Bangka Belitung	(omitted)			0.610	0.129	***	(omitted)			0.651	0.209	***
Province9: Kepulauan Riau	(omitted)			0.910	0.196	***	(omitted)			1.292	0.447	***
Province10: DKI Jakarta	**0.633**	**0.087**	***	**0.644**	**0.063**	***	**0.214**	**0.157**		**0.798**	**0.171**	***
Province11: Jawa Barat	0.520	0.085	***	0.381	0.066	***	0.076	0.141		0.489	0.153	***
Province12: Jawa Tengah	0.243	0.085	***	0.030	0.066		-0.128	0.146		0.315	0.143	**
Province13: D I Yogyakarta	0.053	0.094		-0.019	0.074		-0.010	0.151		0.274	0.161	*
Province14: Jawa Timur	0.253	0.081	***	0.106	0.064	*	0.011	0.134		0.160	0.131	
Province15: Banten	(omitted)			0.580	0.072	***	(omitted)			0.192	0.264	
Province16: Bali	0.187	0.132		0.482	0.103	***	0.071	0.188		0.984	0.238	***
Province17: Kalimantan Barat	(omitted)			1.302	0.448	***	(omitted)			(omitted)		
Province18: Kalimantan Tengah	1.016	0.397	***	1.406	0.274	***	0.429	0.560		0.827	0.491	*

Province19: Kalimantan Selatan	0.405	0.090	***	0.409	0.075	***	0.156	0.128		0.742	0.132	***
Province20: Kalimantan Timur	0.963	0.444	**	0.915	0.139	***		(omitted)		0.614	0.444	
Province21: Sulawesi Selatan	0.111	0.094		0.359	0.077	***	-0.214	0.126	*	0.260	0.120	**
Province22: Sulawesi Tenggara		(omitted)		0.019	0.338			(omitted)		0.714	0.442	
Urban	0.016	0.029		0.120	0.027	***	0.006	0.053		0.273	0.059	***
Constants	6.823	0.148	***	7.957	0.122	***	6.697	0.244	***	6.354	0.340	***
Observation	5326			6728			1060			1391		
The R-squared statistic	0.303			0.263			0.355			0.357		

Source: The author’s calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

3.4.4 Years of Schooling as Education Variable

This part discusses another specification of Mincer wage equation by replacing education level with years of schooling. This is aimed mainly at simplifying the robustness test for these models; besides, the model can estimate the effect of one extra year of schooling on wages.

From all sample estimation (Table 3.25), the result shows that one extra year of schooling increased wages by 12.4 per cent (significant at 1 per cent) in 2000. One extra year of experience also increased the wages by 3.3 per cent, and the effect is diminishing. In 2014, one extra year of schooling and experience affected the wages by 9.9 per cent and 2 per cent, respectively. Gender wise, the result shows that the effect of an additional year of schooling on females is slightly higher than on males in both periods. This result is similar to that of Dumauli's (2015) who finds that the OLS estimates of return to education in Indonesia is around 10 per cent and 12 per cent in 2007; and the return to education of females is higher than males. This result reciprocates with our previous result based on education level. By sector, the return to education in the public sector is higher than in private sector for both periods, this in line with our findings in Section 3.4.3. In short, there is a positive relationship between years of schooling and wages; confirming our findings in the previous section that wages increase in line with the education levels/years of schooling.

Table 3.25: Mincer Equation with Years of Schooling as Education Variable

Log of real hourly wage	All Sample				Female				Male				Private				Public			
	2000		2014		2000		2014		2000		2014		2000		2014		2000		2014	
Years of schooling	0.124	***	0.099	***	0.143	***	0.105	***	0.112	***	0.092	***	0.114	***	0.084	***	0.149	***	0.148	***
	0.005		0.004		0.008		0.007		0.006		0.005		0.005		0.004		0.011		0.012	
Experience	0.033	***	0.020	***	0.020	**	0.021	***	0.045	***	0.022	***	0.032	***	0.006		0.005		0.054	***
	0.005		0.004		0.008		0.007		0.006		0.005		0.005		0.005		0.013		0.013	
Experience squared	-0.001	***	0.000	***	0.000		0.000	*	-0.001	***	0.000	***	-0.001	***	0.000		0.000		-0.001	*
	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	
Sex (1=female)	-0.328	***	-0.266	***	(omitted)		(omitted)		(omitted)		(omitted)		-0.396	***	-0.295	***	-0.049		-0.123	***
	0.025		0.022										0.028		0.024		0.053		0.057	
Married and cohabitate	0.102	***	0.056	*	0.089		-0.079		0.121	***	0.144	***	0.101	***	0.074	**	0.241	**	0.314	***
	0.034		0.033		0.055		0.055		0.042		0.040		0.036		0.034		0.096		0.102	
Other (Separated, divorced and widowed)	-0.051		0.106	*	0.022		0.008		0.043		0.175	*	-0.056		0.105	*	0.180		0.542	***
	0.065		0.059		0.083		0.083		0.120		0.091		0.070		0.062		0.171		0.177	
Religion1: Islam	-0.017		0.067		0.146		0.117		-0.100		0.024		-0.050		0.019		0.045		0.310	
	0.107		0.091		0.173		0.147		0.134		0.114		0.128		0.098		0.178		0.225	
Religion2: Christian/Protestant	0.081		0.244	**	0.353	*	0.352	*	-0.086		0.158		0.100		0.303	**	-0.019		0.223	
	0.121		0.122		0.193		0.196		0.154		0.156		0.145		0.133		0.203		0.295	
Religion3: Catholic	0.047		0.205	*	0.245		0.364	**	-0.081		0.068		0.121		0.229	**	-0.240		0.169	
	0.128		0.106		0.213		0.171		0.157		0.134		0.152		0.115		0.219		0.257	
Religion5: Buddhist	-0.184		0.151		-0.121		0.109		-0.215		0.131		-0.198		0.072		-0.030		1.574	
	0.199		0.302		0.346		0.546		0.241		0.358		0.214		0.304		0.807		0.999	
Ethnicity1: Jawa	-0.043		0.108	***	-0.066		0.136	**	-0.040		0.109	***	-0.051		0.089	**	0.021		0.151	*
	0.047		0.033		0.085		0.056		0.056		0.040		0.054		0.035		0.095		0.089	
Ethnicity2: Sunda	-0.064		-0.133		0.015		(omitted)		-0.132	*	-0.149		-0.061		-0.182		-0.139		(omitted)	
	0.058		0.373		0.102				0.069		0.360		0.065		0.362		0.122			
Ethnicity3: Batak	0.058		0.008		-0.100		0.018		0.127		0.013		0.080		-0.013		0.061		0.000	
	0.094		0.044		0.164		0.077		0.114		0.054		0.106		0.046		0.215		0.140	
Ethnicity4: Betawi	0.005		0.302	*	0.141		0.292		-0.102		0.294		0.018		0.257		-0.312	*	(omitted)	
	0.067		0.167		0.114		0.302		0.081		0.198		0.073		0.163		0.179			
Ethnicity5: Minang	0.052		0.111		0.061		(omitted)		0.002		0.177		0.022		0.229		0.105		(omitted)	

	0.093		0.911		0.160				0.114		0.879		0.106		0.885		0.183			
Ethnicity6: Tiong Hoa	0.109		0.181		0.070		-0.461		0.166		0.417		0.082		0.284		(omitted)		-0.343	
	0.157		0.323		0.265		0.944		0.194		0.333		0.161		0.362				0.693	
Status: fulltime (30 hours a week or more)	-0.646	***	-0.461	***	-0.587	***	-0.352	***	-0.736	***	-0.629	***	-0.696	***	-0.517	***	-0.469	***	-0.324	***
	0.032		0.029		0.049		0.043		0.042		0.040		0.036		0.032		0.068		0.064	
Tenure	0.022	***	0.019	***	0.028	***	0.025	***	0.013	**	0.014	**	0.020	***	0.037	***	0.041	***	-0.001	
	0.005		0.004		0.008		0.007		0.006		0.005		0.005		0.005		0.011		0.012	
Tenure squared	0.000	*	0.000		0.000		0.000		0.000		0.000		0.000	*	-0.001	***	-0.001	**	0.001	
	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	
Sector: private	-0.311	***	-0.289	***	-0.452	***	-0.330	***	-0.234	***	-0.258	***	(omitted)		(omitted)		(omitted)		(omitted)	
	0.038		0.034		0.067		0.055		0.045		0.044									
Industry2: mining and quarrying	0.360	***	0.374	***	1.183	**	0.473		0.314	**	0.431	***	0.322	***	0.369	***	0.787		0.432	
	0.121		0.087		0.501		0.366		0.124		0.089		0.125		0.090		0.552		0.281	
Industry3: manufacturing	0.118	***	0.042		0.140	**	-0.197	**	0.141	***	0.159	***	0.109	**	0.061		0.067		-0.388	
	0.040		0.045		0.070		0.085		0.048		0.053		0.042		0.046		0.149		0.278	
Industry4: electricity, gas and water	0.126		0.133		0.245		-0.337		0.128		0.248	**	0.188		0.229	*	-0.005		-0.430	
	0.143		0.112		0.359		0.483		0.154		0.113		0.183		0.118		0.223		0.322	
Industry5: construction	0.337	***	0.306	***	0.071		0.436	*	0.312	***	0.317	***	0.341	***	0.328	***	0.141		0.079	
	0.049		0.062		0.166		0.225		0.053		0.066		0.051		0.062		0.187		0.504	
Industry6: wholesale, retail, restaurants and hotels	0.070		-0.036		0.124		-0.248	***	0.060		0.084		0.073		-0.019		-0.177		-0.312	
	0.045		0.046		0.077		0.085		0.057		0.055		0.048		0.047		0.205		0.401	
Industry7: transportation, storage, and communications	0.167	***	0.104		-0.186		0.140		0.162	***	0.158	**	0.170	***	0.109		0.199		0.659	*
	0.057		0.072		0.222		0.230		0.060		0.075		0.062		0.072		0.155		0.373	
Industry8: finance, insurance, real estate and business services	0.434	***	0.224	***	0.538	***	0.125		0.368	***	0.280	***	0.490	***	0.253	***	0.296		0.153	
	0.094		0.057		0.157		0.112		0.115		0.064		0.105		0.058		0.202		0.205	
Industry9: social services	0.114	***	-0.136	***	0.126	*	-0.300	***	0.158	***	-0.049		0.091	**	-0.124	***	0.138		-0.167	
	0.038		0.044		0.070		0.082		0.046		0.052		0.042		0.046		0.100		0.142	
Firm size2: 5-19 people	0.183	***	0.067	**	0.336	***	0.013		0.106	***	0.084	**	0.174	***	0.061	**	0.170	**	0.116	
	0.027		0.029		0.049		0.048		0.033		0.036		0.030		0.030		0.072		0.103	

Firm size3: 20-99 people	0.249	***	0.243	***	0.426	***	0.272	***	0.147	***	0.205	***	0.266	***	0.233	***	0.194	**	0.301	***
	0.032		0.031		0.056		0.053		0.039		0.038		0.035		0.033		0.077		0.104	
Firm size4: >= 100 people	0.339	***	0.513	***	0.585	***	0.620	***	0.216	***	0.446	***	0.409	***	0.526	***	0.129		0.556	***
	0.043		0.035		0.070		0.060		0.054		0.042		0.047		0.036		0.107		0.116	
Province1: Aceh	(omitted)		1.348		(omitted)		(omitted)		(omitted)		1.345		(omitted)		1.191		(omitted)		(omitted)	
			0.917								0.886				0.892					
Province2: Sumatera Utara	0.361	***	0.417	***	0.625	***	0.444	***	0.277	***	0.412	***	0.462	***	0.365	***	0.065		0.628	***
	0.086		0.061		0.149		0.103		0.103		0.075		0.096		0.070		0.204		0.133	
Province3: Sumatera Barat	0.251	**	0.500	***	0.211		0.728	***	0.306	**	0.353	***	0.312	**	0.405	***	0.086		0.629	***
	0.111		0.061		0.192		0.100		0.133		0.078		0.127		0.073		0.210		0.118	
Province4: Riau	(omitted)		0.796	***	(omitted)		0.985	***	(omitted)		0.758	***	(omitted)		0.653	***	(omitted)		1.275	***
			0.126				0.272				0.140				0.135				0.332	
Province5: Jambi	(omitted)		0.579	**	(omitted)		0.695		(omitted)		0.432		(omitted)		0.504	*	(omitted)		(omitted)	
			0.267				0.552				0.299				0.261					
Province6: Sumatera Selatan	0.198	**	0.362	***	0.240		0.512	***	0.185	**	0.287	***	0.264	***	0.317	***	0.050		0.340	**
	0.080		0.068		0.147		0.122		0.094		0.080		0.091		0.076		0.158		0.151	
Province7: Lampung	0.150		0.166	**	0.068		0.103		0.166		0.176	**	0.202	*	0.096		0.112		0.217	
	0.092		0.075		0.176		0.135		0.108		0.090		0.103		0.082		0.213		0.195	
Province8: Kepulauan Bangka Belitung	(omitted)		0.570	***	(omitted)		0.695	***	(omitted)		0.497	***	(omitted)		0.585	***	(omitted)		0.602	***
			0.110				0.184				0.135				0.129				0.208	
Province9: Kepulauan Riau	1.037	***	0.942	***	1.279	***	1.217	***	0.931	***	0.886	***	1.176	***	0.891	***	0.052		1.275	***
	0.156		0.181		0.310		0.393		0.179		0.200		0.168		0.195		0.470		0.445	
Province10: DKI Jakarta	0.543	***	0.685	***	0.743	***	0.802	***	0.467	***	0.617	***	0.637	***	0.637	***	0.220		0.797	***
	0.076		0.057		0.133		0.099		0.093		0.069		0.087		0.063		0.157		0.170	
Province11: Jawa Barat	0.419	***	0.430	***	0.526	***	0.527	***	0.395	***	0.376	***	0.526	***	0.381	***	0.080		0.458	***
	0.074		0.060		0.131		0.102		0.088		0.073		0.085		0.066		0.141		0.152	
Province12: Jawa Tengah	0.165	**	0.113	*	0.221	*	0.114		0.150	*	0.117		0.250	***	0.037		-0.123		0.305	**
	0.075		0.059		0.132		0.099		0.089		0.073		0.086		0.066		0.146		0.143	
Province13: D I Yogyakarta	0.020		0.059		0.109		0.172		0.015		-0.004		0.049		-0.025		-0.004		0.262	
	0.081		0.067		0.140		0.112		0.099		0.083		0.094		0.074		0.151		0.161	
Province14: Jawa Timur	0.181	**	0.136	**	0.347	***	0.176	*	0.119		0.123	*	0.250	***	0.106	*	0.020		0.158	

	0.070		0.057		0.126		0.096		0.084		0.069		0.081		0.064		0.134		0.131	
Province15: Banten	(omitted)		0.628	***	(omitted)		0.885	***			0.447	***	(omitted)		0.578	***	(omitted)		0.155	
			0.067				0.111				0.083				0.072				0.263	
Province16: Bali	0.137		0.602	***	0.568	***	0.846	***	-0.051		0.458	***	0.179		0.482	***	0.071		0.978	***
	0.111		0.094		0.185		0.155		0.138		0.118		0.132		0.103		0.188		0.237	
Province18: Kalimantan Barat	(omitted)		1.373	***	(omitted)		1.641	**	(omitted)		1.113	*	(omitted)		1.340	***	(omitted)		(omitted)	
			0.459				0.675				0.625				0.447					
Province19: Kalimantan Tengah	0.736	**	1.318	***	1.422	**	0.930	**	0.578		1.485	***	0.959	**	1.414	***	0.457		0.840	*
	0.332		0.240		0.617		0.430		0.389		0.285		0.398		0.274		0.559		0.489	
Province20: Kalimantan Selatan	0.331	***	0.514	***	0.652	***	0.440	***	0.219	**	0.538	***	0.400	***	0.401	***	0.156		0.734	***
	0.076		0.065		0.142		0.112		0.088		0.078		0.090		0.075		0.128		0.132	
Province21: Kalimantan Timur	0.902	**	0.922	***	0.481		0.810	***	1.039	**	0.990	***	0.941	**	0.899	***	(omitted)		0.600	
	0.438		0.134		0.878		0.236		0.500		0.160		0.445		0.139				0.443	
Province22: Sulawesi Selatan	0.006		0.319	***	0.209		0.175	*	-0.094		0.442	***	0.104		0.356	***	-0.213	*	0.257	**
	0.077		0.065		0.135		0.106		0.093		0.081		0.094		0.077		0.125		0.120	
Province23: Sulawesi Tenggara	(omitted)		0.200		(omitted)		-0.289		(omitted)		0.324		(omitted)		0.021		(omitted)		0.683	
			0.266				0.549				0.297				0.338				0.440	
Urban	-0.004		0.140	***	-0.040		0.177	***	0.023		0.121	***	0.013		0.116	***	0.004		0.261	***
	0.026		0.024		0.047		0.041		0.031		0.030		0.029		0.027		0.053		0.059	
Constants	6.286	***	7.842	***	5.496	***	7.595	***	6.508	***	7.958	***	6.109	***	7.404	***	5.770	***	5.374	***
	0.140		0.163		0.237		0.264		0.172		0.207		0.155		0.129		0.264		0.333	
Observation	6386		8119		2170		3173		4216		4946		5326		6728		1060		1391	
R Square	0.36		0.30		0.46		0.33		0.30		0.28		0.29		0.26		0.33		0.34	

Source: The author’s calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

3.4.5 The Effects of Education Expansion in Indonesia

To estimate the change in return to education, the present study calculates the percentage of the change and runs a Seemingly Unrelated Regression (SUR)[24] for both periods, implemented via Stata's -suest- command, followed by the Wald test with the null hypothesis that the coefficients of the return to education are common to both models and identical ($\beta_{1,t-1} = \beta_{1,t}$), for example: the coefficient of return to junior high school in 2000 is equal to the coefficient of return to junior high school in 2014. This is aimed at testing the equality of coefficients across two models or Mincer wage equation for 2000 and 2014. On the other hand, this become the limitation of this study, i.e.: restrictiveness because the present study imposes a common coefficient of return to education across the years.

From the test result (Table 3.26), the hypothesis is rejected at 1 per cent significance level. Thus, the test confirms that there was a decrease in the returns to each education level in both periods. From 2000 to 2014, the returns of each education level decreased. In the basic regression, the highest decrease was the return of university (relative to primary school and below), at -0.329 percentage point. The coefficient of experience decreased slightly from 5.9 per cent in 2000 to 4.4 per cent in 2014. Experience square coefficient was still the same, implying that there is a diminishing return of experience on wages.

In full regression, the result shows a similar pattern of decline as what Gropello and Sakellariou (2010) find (Table 3.4); the decline of return to university (-0.235 percentage point) was higher than the decrease of return to senior high school (-0.157 percentage point), and the hypothesis of the Wald test is rejected, or the coefficients are not identical and there is a decline in the return to education. Meanwhile, the decline of return to junior high school was around –0.064 percentage point, however, the Wald test's hypothesis cannot be rejected.

[24] The SUR is a generalization of a linear regression model that consists of several regression equations; each equation has its own dependent variable and potentially different sets of exogenous explanatory variables. And, each equation is a valid linear regression on its own and can be estimated separately. The errors are assumed to be correlated across the equations since those equations have the same unobservable variables such as ability. The error of each equation may have its own variance. Each equation is assumed correlating with the others in the same time period.

Similar result is also obtained for the estimation based on gender (Table 3.27) and private sector (Table 3.28); the return of senior high school and university decreased, except for the junior high school coefficient. Wald test results also confirm that return for males decreases (at 10 per cent significance level). A possible reason for the contrasting results of the junior high school is because the GER of primary and junior high school is already relatively high (Figure 2.2); subsequently, the expansion of education affects more significantly on higher education levels (senior high school and university in particular).

In contrast, the public sector yields different results; the hypotheses of proportionality of the coefficients of the two models cannot be rejected. The changes are insignificant. This is possibly due to the wage rigidity in the public sector as the sector has its own wage mechanism that is not based on the market's mechanism, as explained in Chapter 2. It is worth noting that public and private sectors in developing countries are non-competing, as asserted by Fields (2011), in which individuals belong to one labour market segment or another, and they cannot or will not switch from one to another. Other possible explanations are firstly, the non-transparent remuneration, as McLeod and Macintyre (2007) assert that civil servants in Indonesia receive a range of supplementary allowances in addition to their basic salaries. Some of these allowances (such as payments for attending meetings) border on the absurd, because they amount to payments simply for doing one's job. These additional allowances make civil servants' overall remuneration non-transparent, and certainly make their take-home pay considerably higher than what the published basic pay level would indicate. Secondly, following the bureaucracy reform initiative implemented in the Ministry of Finance in 2007, there is an additional allowance that substantially increases civil servants' take-home pay, particularly for workers in several ministries such as the Ministry of Finance (Tjiptoherijanto, 2015).

Table 3.26: The Change of Main Variables, All Individuals

Log of real hourly wage	All Sample						
	Basic Regression (coefficients)			Full Regression (coefficients)			SUR
	2000	2014	Change	2000	2014	Change	Wald Test
Junior High School	0.427	0.326	-0.101	0.285	0.221	-0.064	
Senior High School	0.899	0.755	-0.144	0.637	0.480	-0.157	***
University	1.597	1.268	-0.329	1.193	0.958	-0.235	***
Experience	0.059	0.044	-0.015	0.032	0.022	-0.01	
Experience squared	- 0.001	-0.001		- 0.001	0.000		
Constants	6.369	7.649		7.040	8.471		
Observation	6386	8119		6386	8119		
The R-squared statistic	0.26	0.15		0.37	0.30		

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table 3.27: The Change of Main Variables, by Gender

Log of real hourly wage	Female				Male			
	2000	2014	Change	Wald Test	2000	2014	Change	Wald Test
Junior High School	0.251	0.305	0.054		0.281	0.174	-0.107	*
Senior High School	0.770	0.583	-0.187	**	0.572	0.421	-0.151	**
University	1.311	1.015	-0.296	***	1.101	0.897	-0.204	***
Constants	6.438	8.258			7.156	8.557		
Observation	2170	3173			4216	4946		
The R-squared statistic	0.471	0.333			0.314	0.285		

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table 3.28: The Change of Main Variables, by Sector

Log of real hourly wage	Private				Public			
	2000	2014	Change	Wald Test	2000	2014	Change	Wald Test
Junior High School	0.259	0.178	-0.081		0.335	0.155	-0.18	
Senior High School	0.577	0.412	-0.165	***	0.825	0.664	-0.161	
University	1.177	0.824	-0.353	***	1.290	1.317	0.0267	
Constants	6.823	7.968			6.697	6.386		
Observation	5326	6728			1060	1391		
The R-squared statistic	0.303	0.264			0.355	0.359		

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

This decline is also in line with the global trend, for instance, Gropello *et al.* (2011) find a sharp decline in returns during the past few decades, reflecting the sharp rise in schooling levels worldwide; the world population aged 15 and above was estimated to have an average of 8 years of schooling in 2010, having increased steadily from just over 5 years in 1980. In the meantime, the returns to schooling have declined significantly since the 1980s, when they were above 13 percent, to just over 9 percent in recent years.

The highest decrease in return to education is the return for males with junior high school qualifications (relative to primary school or below), probably due to more workers having higher education attainment in 2014/2015. A possible explanation for the proliferation of educated workers is the success of the 9-year (primary and junior high school) compulsory education program; thus, the expanding supply side. Correspondingly, on the demand side, more companies prefer higher qualifications, for example: an increase in retail sector causes supermarkets and mega-malls to replace the traditional market. This, subsequently, decreases the demand for unskilled labour within these sectors, particularly the demand for unpaid family workers, and *vice versa* increases the demand for workers with higher education attainment (Allen, 2016), at least with senior high school qualifications. As a result, the supply of junior high school graduates outstrips the demand and return to education of junior high school decreases more substantially than other levels, as confirmed by the data of job seekers and vacancies (Figure 2.13).

Similarly, return to senior high school decreased from 63.7 per cent in 2000 to 48 per cent in 2014, or around -24.6 per cent (in full regression result). Based on gender, males could have more contribution to this decrease of -26.40 per cent. This is probably because of employment outcomes in manufacturing, which would have favoured the males, have been modest between 2005 and 2014 period (Allen, 2016). In similar vein, Purnastuti *et al.* (2013) and Gropello and Sakellariou (2010) note how most of the declines occurring in the senior high school level is associated to vocational senior high schools, because specialised skills may lose value more rapidly in a dynamic labour market.

By sector, higher decrease has been experienced by the private sector. A possible reason is due to the more competitive nature of the private sector; as Amiti (2011) asserts that wages in Indonesia are largely determined by the market, with the exception of minimum wages which are set by provincial governments. Likewise, Psacharopolous (1979, in Chevalier *et al.*, 2002) elaborates on a similar distinction between the private and the public sectors and argues that wages could exceed productivity in the public sector but not in the more competitive private sector. As such, the lack of competition in the public sector allows higher returns to education in that sector.

The result also shows that the return to university decreased in all individuals, around -19.7 per cent, from 119.3 per cent in 2000 to 95.8 per cent in 2014. Figure 2.13 shows that the number of vacancies for university is still slightly higher than the number of job seekers, but the data were taken in 2011, and unfortunately there is no updated data for 2014, hence there is no information whether or not the trend has changed. A possible explanation is the mismatch between the supply and demand of university graduates, which could be in terms of quantity, quality, subject (engineers) or skill.

There are several empirical studies reveal that there is a mismatch between labour demand and supply in terms of quality and quantity in Indonesia. Firstly, Boston Consulting Group (2013) predicts that there would be a 40 per cent to 60 per cent gap between the demand and supply for middle management jobs in 2020. Meanwhile, the shortage of entry level problems will worsen quickly. Consequently, top companies cannot fill about one-half of their entry-level positions with fully qualified candidates. Boston Consulting Group also predicts that service sector expansion in the future will increase administrative and managerial positions from 36 per cent in 2013 to 55 per cent in 2020. Meanwhile, Indonesia already produces fewer graduates for those positions than the current demand. Moreover, Dumauli (2015) also reveals that high skilled jobs are not in demand in the country; as a result, the return to education decreases. Besides that, the low return to education in Indonesia may be induced by the low quality of the education system. In other words, there is an excess of supply of university graduates, which in turn reduces the wage of university graduates.

Regarding education quality, Beatty *et al*. (2018) assert that the learning level declines during 2000 and 2014 periods; the downward trend occurs for a long period and is expected to persist without any interventions, as explained in Chapter 2, Figure 2.1. PISA result also confirms that education quality in Indonesia is relatively low as the country only performs below the 25th percentile of the OECD average (World Bank, 2018).
In terms of sector, it seems that the market determines only the wages in the private sector, while the public sector has its own mechanism, as explained in Section 3.4.3. However, this result is in contrast to Purnastuti *et al* (2013) who find that the profitability of university degrees in Indonesia increases between 1993 and 2007/08 for both males and females.

In terms of years of schooling (Table 3.29), the result also shows a decline in return to education in Indonesia. The coefficient of years of schooling for all individual in 2000 was 12.4 per cent, which then decreased to 9.9 per cent in 2014. Although the return to education of females was higher than that of males, females experienced more significant decrease in the return of schooling (from 14.3 per cent to 10.5 per cent) than male (11.2 per cent to 9.2 per cent). A possible explanation is that young females in urban areas have increased their labour participation in recent years (Schaner and Das, 2016). Finally, the return to private sector experienced a significant decline from 11.4 per cent to 8.4 per cent during the period, which is in line with previous analysis that competition in the private sector may lead to this significant decrease. The hypothesis of Wald test for public sector is also rejected, similar to the education level analysis.

Table 3.29: Wald Test for Years of Schooling as the Education Variable

Years of schooling	2000		2014		Change	Wald Test (SUR)
All Sample	0.124	***	0.099	***	-20.0	***
Std. Err.	0.005		0.004			
Observation	6386		8119			
R Square	0.36		0.30			
Female	0.143	***	0.105	***	-26.3	***
Std. Err.	0.008		0.007			
Observation	2170		3173			
R Square	0.46		0.33			
Male	0.112	***	0.092	***	-17.6	**
Std. Err.	0.006		0.005			
Observation	4216		4946			
R Square	0.30		0.28			
Private	0.114	***	0.084	***	-26.3	***
Std. Err.	0.005		0.004			
Observation	5326		6728			
R Square	0.29		0.26			
Public	0.149	***	0.148	***	-0.7	
Std. Err.	0.011		0.012			
Observation	1060		1391			
R Square	0.33		0.34			

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

In general, the estimation results show that the return to each level of education decreases between 2000 and 2014, as per Gropello and Sakellariou (2010). In addition, the return to one additional year of schooling also confirms this decrease. A possible explanation for these declines is that Indonesia is in the third stage of education expansion, hence the shifting of the supply side. However, the demand side in the waged sector has not yet adjusted optimally; in other words, changes in the labour supply outstrip the changes in the labour demand. Similarly, Allen (2016) asserts that slow job growth has been a persistent challenge for Indonesia, due to the combined slower rates of economic growth and job creation which have limited the expansion of quality jobs and slowed the pace of structural transformation in the country. As such, the relative employment rate increases while the relative wage decreases, as explained by the supply and demand framework in the literature review.

3.4.6 Control Variables

In terms of the control variables, some control variables affect the wage both in 2000 and 2014 period, such as: sex, job status, tenure, sector and some industries as well as some dummies of province. The present study's finding for those control variables are consistent with previous empirical studies, as mentioned in Section 3.2.2, with the exception of job status variable that need further investigation.

The result shows that coefficient of sex is negative and significant at 1 per cent in all sample estimation (Table 3.20), including OLS with years of schooling. Even for the results by sector, gender affects wages in both public and private sectors, implying that males receive higher hourly wages than females, as in line with Becker's employer taste model of discrimination and the statistical discrimination theory. Becker's model is based on the idea that some workers, employers or customers do not want to work with or come into contact with members of other racial groups or with women[25] (Becker, 1971). Meanwhile, the statistical discrimination theory asserts that inequality may exist and persist between demographic groups even when economic agents (consumers, workers,

[25] This is to say that discrimination can persist only if there are factors which limit the amount of competition in the labour market or in the product market. If these markets are competitive, the increased profitability of non-discriminating firms compared to the discriminating ones will encourage non-discriminators to enter the market. This will put downward pressure on the price level and eventually force the higher-cost discriminating firms out of business (Becker, 1971).

employers, *etc*.) are rational and not prejudiced. This result is also similar to those of Comola and de Mello's (2013), Purnastuti's *et al* (2015), and Dong's (2016).

Married and cohabitate status coefficient was positive and significant at 1 per cent (different from single status) in 2000 but turned insignificant in 2014 (Table 3.20). The other marital dummy was insignificant. In regression by gender, the result shows that the married status affects males in both periods (Table 3.22). In regression by sector, married and cohabitate status had a positive and significant effect on wages in both public and private sectors (Table 3.23). A possible explanation is that family allowances are part of wage, particularly for the public sector (The World Bank, 2000). Subsequently, married workers can have higher wages than single workers whereas in the literature, the expected sign of marital status coefficient is *ex ante* unclear. This could have positive relationship, since married individuals might have better health, thus having more energy and higher productivity (Guner *et al.,* 2018). Furthermore, marriage can also generate efficiencies through specialisation and the division of labour where tasks are divided between spouses, thus freeing up time (Baker and Jacobsen, 2007). On the contrary, the married status could have a negative relationship, which might occur when individuals have less time available for work because of family commitments (Mishra and Smyth, 2013).

Turning to religion, most religion dummies were not significantly different, relative to Hindu religion, except for Catholic and Christianity/Protestant in 2014 (positive and significant at 5 per cent) as shown in Table 3.22. The results should be interpreted cautiously; this (probably) does not mean that certain religion has higher wages. Instead, this is probably because of the average education of the believers of certain religion is higher than the other, for example: most Christian/Protestant and Catholic individuals have higher education, more than 80 per cent has (at least) senior high school and university qualifications.

Furthermore, all ethnicity dummies were statistically insignificant, relative to other ethnicities in 2000. The result slightly changed in 2014; several majority ethnicities have a positive and significant effect on wages, for instance Jawa (significant at 1 per cent) and Betawi (significant at 10 per cent). Estimations based on gender and sector show that the Jawa ethnicity (the Javanese) leads to higher wages for both males and females, and in both public and private sectors. It seems that the majority ethnicity still has an advantage.

Regarding work related and firm size variables; the first variable, tenure, had a positive and significant coefficient at 1 per cent in most specifications, as in line with Chevalier *et al.* (2002) and Purnastuti *et al.* (2015), except in the public sector in 2014. Tenure squares were insignificant in most regressions, except for regression by sector, which may imply that tenure does not have a diminishing effect on wages in general and by gender. A possible argument is that regular employees, the youth, women and those with lower levels of educational attainment tend to have shorter periods of job tenure than other workers, suggesting that these groups are more likely to work on short-term contracts than on permanent ones, due to the segmentation in the labour market (Allen, 2016).

In terms of sector, the private sector yielded a negative and significant coefficient at 1 per cent for each specification (all individuals and by gender) in both periods, implying that wages in the private sector was less than wages in the public sector. This is in line with the report from World Bank (2000) that civil service salaries have increased significantly in real terms and at a much faster rate than those in the private sector; thus, the average government worker earned more, not less, than her/his private sector counterpart.

Additionally, most of the industry dummies show a positive and significant difference, relative to agriculture sector at 5 per cent. While in 2014, more sectors were insignificant, such as: (1) manufacturing; (2) electricity, gas and water; (3) wholesale, retail, restaurant and hotels; and (4) transportation, storage, and communication industry. Possible explanations for this insignificant difference are: the sample is from the waged sector (public and private companies), some of them are formal sector and generally the minimum wage laws cover the formal sector (ASEAN, 2013); moreover, the labour intensive manufacturing sectors have experienced a decline (Aswicahyono *et al.*, 2011), as have the agriculture sectors.

Regarding the full-time dummy variable, the result shows that all coefficients are negative and significant. However, cautious interpretation is needed, as this (probably) does not mean that full-time workers earn less than part-time workers. In the detailed summary statistic of the sample from 2014, workers with less than 30 hours a week were relatively more educated (40 per cent are university graduates and 25 per cent are high school graduates); had more experience (the mean value being 16.7 year); and mostly work in social services (62.4 per cent)[26]. This is related to one of the (non-financial) incentives of being public sector workers in Indonesia being the flexible working hours (UNDP, 2014). In the private sector, some companies in Indonesia have already offered flexible working hours arrangements, such as Kraft foods which offers extended leave (sabbatical) and Dow Chemical which offers adjusted working hours to care for a child with special needs/or elders, telecommuting, and compressed work week. Offering flexible working hour is also driven by the traffic congestions that appear to be a particularly challenging problem in Indonesia's urban areas. Furthermore, Kraft estimates that 100 per cent of their employees use their flexible time at least occasionally, while Dow estimates that 15 per cent of their workforce utilises this option (Anell and Hartmann, 2007).

The last control variable in work related and firm size category is the firm size. The result shows that all of the coefficients were positive and significantly different (at 1 per cent) from firms with only 1-5 workers (small firms). In general, this implies that the bigger the firms are; the higher the wages that the workers can earn. This is in line with Dhanani and Islam (2004) who assert the labour productivity varied substantially according to the size of the manufacturing units. An exception for the public sector is that the central government determines the wages and the changes, if any, by government regulation.

With regards to region/residence as control variables, most provinces show a positive and significant difference with the benchmark province (Nusa Tenggara Barat/NTB), except for DI Yogyakarta (located in Jawa Island, as the centre of the economy). A possible explanation is that the living cost and the minimum wage in DI Yogyakarta are very low. Meanwhile, the capital province (DKI Jakarta) has a higher wage than NTB. The present study has omitted some provinces (besides NTB): Aceh, Sumatera Selatan, and Sulawesi Tenggara in 2000; and Papua Barat in 2014. This is due to the absence of individuals from those provinces in those periods, or that no one in those provinces past our restriction, particularly the employment restriction, or there is a missing variable in the sample such as missing wages, thus rendering those individuals unanalysable.

The last control variable is the urban/rural dummy (Urban=1). The result shows that this dummy was insignificant in 2000; but turned positive and significant at 1 per cent in 2014.

[26] On the other hand, only 23.5 per cent workers who worked more than 30 hour a week have university qualifications; their mean of experience was less than the other group (16 years); they worked mostly in social services (31 per cent), manufacturing (24.5 per cent) and wholesale, retail, restaurants and hotels (19 per cent).

A possible explanation why urban and rural areas were insignificantly different in 2000 is because rural wages rose in most Asian countries in the early 2000s, including some provinces of Indonesia. The potential drivers of rural wages are the incline in agricultural labour productivity, growth of manufacturing output and the increase in rural working populations (Wiggins and Keats, 2014). In 2014, there was a significant difference between the urban and rural as urban areas affect higher hourly wages in all individuals, by gender and by sector, as is in line with Comola and de Mello (2011).

3.4.7 Robustness Test

This part will address the issue of Mincer wage equation limitation, *i.e*. the endogeneity and sample selection bias. This robustness test is aimed to investigate the decrease that also occurs in IV and Heckman model, which is in line with the OLS results. Besides the instrument used in this part, other possible instruments in both years based on literature are provided in Appendix VII.

Endogeneity

The Instrumental Variable (IV) is used to gauge the role of omitted variables (ability bias) in the OLS estimates of the return to schooling in the Indonesian labour market. This part discusses the robustness test for all individuals, a similar model to the OLS. Sector and gender analysis is attached in Appendix VIII for further information.

Following previous studies (see Duflo, 2001; Comola and de Mello, 2010; Purnastuti *et al.*, 2015) which find that there is a strong correlation between education policies and the increase of enrolment (supply side shifting) as well as policy instruments are used in this study. There are three different policy instruments used; the school construction program (INPRES), the 6-year compulsory education program (CSAL1), and the 9-year compulsory education program (CSAL2), as explained in Chapter 2. However, this study will not use all of the instruments in each model, considering that those policies have the similar aim to increase primary school participation.

The main advantages of using policy instruments are that it is relatively easy to construct by defining the dummy variables and it retains the number of observations in our sample (Appendix VII). The idea of policy instruments is that an individual who was born before the policies were implemented would have a lower level of education compared to the people born later, who would be affected by the policies.

Despite the policy instruments' advantages, Purnastuti *et al.,* (2015) highlight some limitations; firstly, variables for compulsory schooling laws may confound changes due to the cohort effects with these laws; and secondly, the accessibility or availability of schooling instruments may be sensitive to the relative size of the returns to schooling of the groups most affected by the changed conditions reflected in the instrument (a heterogeneity in returns to education argument).

The break points of those policies are 1967, 1977 and 1987; the year of 1974 refers to the year when the primary school buildings were completely constructed under the INPRES program and the age of 7 was the official age to start primary education. The dummy variable for the INPRES program therefore has a value of 1 for individuals born after 1967 and zero for all other individuals; or, INPRES=1 if individual was born in 1967 and later, otherwise 0. CSAL1=1 if an individual was born in 1977 and later, otherwise 0; and CSAL2=1 if an individual was born in 1987 and later. The sample's age range is 16-55 years old (referring to the age restriction), and mostly was born before the year of 1987; thus, all individuals in 2000 were not affected by CSAL2. *Vice versa*, 90 per cent of respondents in 2014 were affected by INPRES, since all of them were born after 1967. In additions, the distribution of year of schooling and birth cohort is provided in Appendix VIII, Table VIII.1. For the IV model of the year 2000, the present study does not consider using CSAL1 as the instruments because the sample which is affected by the policy are aged between 16-23 years old; and 23 years of age is too young in this case, as most Indonesians only finish university at the age of 22-23 years old. Subsequently, INPRES will be the only instruments of that year.

Furthermore, CSAL1 and/or CSAL2 are selected as the instruments for 2014 (considering that CSAL2 substitutes for and is the expansion of CSAL1). In contrast to the year 2000, this study does not consider INPRES because most individuals are aged 16-44 years old, and most of them are affected by the policy (around 90 per cent of the sample). Moreover, INPRES is already replaced by CSAL1 in recent years. In addition, this study will estimate two IV models for 2014, the first one being both CSALs (CSAL1 and CSAL2) and the second being only CSAL1. The distributions of individuals based on instruments, education level, and age are provided in Appendix VIII.

Table 3.30 presents the estimates of Mincer wage equation using policy instruments for the year 2000. The reduced-form-schooling equation has a relatively small R squared (around 10 per cent), while F-statistic in first stage estimation is relatively high, at 701.24, and significant at 1 per cent, suggesting that the instrument is not weak. The coefficient of years of schooling is 0.108 (the second stage), significant at 1 per cent, which is slightly lower than its coefficient in the OLS model (0.124). However, INPRES has a significant and negative coefficient. The 2014 model also shows a similar result; both of CSAL1 and CSAL2 instruments are negative and significant effects on years of schooling (endogenous variable), this could be due to failure to allow for time trend/time dummies or in this context is the instruments, one possible reasons is there is multicollinearity between the instruments and experience variables, both of variable could be highly linearly related. F-tests are relatively high, more than 1000 (Table 3.31). Likewise, the coefficients of years of schooling are 0.066 (Model 1) and 0.088 (Model 2), which is also slightly lower than coefficient of the OLS model (0.099).

This result's interpretation is inconsistent with the idea of policy instruments (an individual born earlier or before the policy is implemented would have a lower level of education as compared to the people born later who would be affected by the policies). It is also contradictory with the aim of the policies to increase the school participation rate. A possible reason for this unexpected sign is the multicollinearity between the policies and experience, since we define experience from potential experience (age – years of schooling – starting school age), and policy instruments are also defined by age, as explained previously. According to Aronow and Carnegie (2013) IV estimation is commonly used, but it is subject to an often-overlooked limitation. Wald (1940) in Aronow and Carnegie (2013) revealed many researchers intend to estimate the average treatment effect (ATE) for the entire population interest, but IV estimation only covers the local average treatment effect (LATE) or the ATE for the subpopulation that is influenced by the IV. When treatment effect are heterogenous across units, the LATE and the ATE may take on different values, this potentially causing complication in the interpretation, as well as makes comparison difficult.

This study also tried to estimate by using alternative models by removing experience and experience squared (Appendix VIII), then testing the instruments with hypothesis: instruments are weak. The result shows that the partial R-squared are relatively low, but the F-tests have p-value of less than 5 per cent, which indicates that the instruments are not weak. In contrast, with the models including experiences, the policy instruments now have a positive sign and a significant effect, which implies that individuals who are affected by the policies would have an extra year of schooling compared with individuals who are not, except for the 9-year compulsory education (CSAL2) in 2014, which is insignificant. A possible argument is that the age distribution of individuals who are affected by CSAL2, which is 16-27 years old (29.6 per cent of total sample), is too young compared to the entire individuals' age distribution.

The results of the endogeneity test for Model 1 of 2014 show a very small p-value (less than 5 per cent), suggesting that we can reject H0; thus, years of school variable is an endogenous variable, or the endogeneity problem occurs in that model. Meanwhile, for Model 1 of 2000 and Model 2 of 2014, the p-values are insignificant, or the endogeneity problem may not occur. This agrees with Duflo (2001) and Comola and de Mello (2010) who report that there is little evidence of ability bias in the OLS estimates of the return to schooling in Indonesia, and that the OLS estimates are not likely to be biased upwards.

The Anderson canonical correlation LM statistic can check rank condition for identification. The result shows that the excluded instruments are correlated with the endogenous regressor and the equation is thus identified. Meanwhile, Sargan statistic (overidentification test of all instruments) suggests that the estimations are exactly identified, since we have one endogenous variable and one instrument, except for model 1 of 2014 where we have two instruments for one regressor. Thus, we have an overidentification in Model 1 of 2014. In spite of the instruments' weaknesses, often-overlooked limitation, and complication in the interpretation (Aronow and Carnegie, 2013), the IV result confirms that there is a decline of year schooling coefficient from 0.108 in 2000 to 0.066 (Model 1) and 0.088 (Model 2) in 2014.

Table 3.30: IV Specification with Years of Schooling in 2000, All Individuals

2000: Log of real hourly wage	1st Stage Coef.	Std. Err.	P>t	2nd Stage Coef.	Std. Err.	P>z
Years of schooling				0.108	0.015	***
Experience	-0.403	0.013	***	0.029	0.006	***
Experience squared	0.003	0.000	***	-0.001	0.000	***
…	…	…		…	…	
INPRES	-2.788	0.105	***			
Constant	15.204	0.358	***	6.467	0.215	***
Observation	6386			6386		
R Squared	0.10			0.37		
Test Results on Instruments						
F-Test				701.24		***
Underidentification test (Anderson canon. corr. LM statistic):				636.17		***
Sargan statistic (overidentification test of all instruments):				Exactly identified		
Endogeneity test of endogenous regressors:				1.23		

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table 3.31: IV Specification with Years of Schooling in 2014, All Individuals

2014: Log of real hourly wage	Model 1: 1st Stage Coef.	Std. Err.	P>t	Model 1: Second Stage Coef.	Std. Err.	P>z	Model 2: 1st Stage Coef.	Std. Err.	P>t	Model 2: 2nd Stage Coef.	Std. Err.	P>z
Years of schooling				0.066	0.008	***				0.088	0.011	***
Experience	-0.547	0.013	***	0.015	0.004	***	-0.231	0.011	***	0.019	0.005	***
Experience squared	0.004	0.000	***	0.000	0.000	***	-0.001	0.000	***	0.000	0.000	***
…	…	…		…	…		…	…		…	…	
CSAL1	-3.950	0.088	***				-3.303	0.095	***			
CSAL2	-3.652	0.092	***									
Constant	25.622	0.366	***	8.447	0.202	***	20.861	0.379	***	8.052	0.248	***
Observation	8119		***	8119			8119			8119		
The R-squared statistic	0.27			0.30			0.13			0.30		
Test Results on Instruments												
F-Test				1521.18		***				1216.65		***
Underidentification test (Anderson canon. corr. LM statistic):				2224.24		***				1064.36		***
Sargan statistic (overidentification test of all instruments):				7.36		***				Exactly identified		
Endogeneity test of endogenous regressor:				24.42		***				1.03		

Source: The author’s calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Furthermore, the present study also performs IV based on conventional instruments (father and mother education), but the result is very sensitive to the sample size, and using conventional instruments reduces sample size up to around half of the sample of each year, thus the results of both the OLS and the IV changes are significantly different to the models in this research.

Another alternative instrument is smoking (as one of bad habit variables), and the result is provided in Appendix IX. The idea of using bad habit as an instrument is because more educated people have better health and better health habits and completed years of formal schooling is the most important correlation of good health (Grossman, 2008). Smoking variable is defined in question KM01a in IFLS questionnaires "Have you ever chewed tobacco, smoked a pipe, smoked self-rolled cigarettes, or smoked cigarettes/cigars?". The dummy is 1 if 'Yes' and 0 otherwise. The result shows that smoking dummies have a negative correlation with years of schooling, as expected, and are significant at 1 per cent. The coefficient of years of schooling of the IV model is slightly lower than the OLS' coefficient for the same specification. For example: the coefficient of years of schooling of the OLS is 0.120 and the coefficient of the IV is 0.102 in the year 2000; likewise, the coefficient of IV in 2014 is 0.088, slightly lower than the OLS' of 0.099. The F-test for the year 2000 is 35 and for 2014 is 71, indicating that the instrument is not weak. Comparing both periods, the IV results also confirm that there is a decrease of year schooling variable from 0.102 in the year 2000 to 0.088 (around -13.7 per cent) in 2014. Additionally, the endogeneity tests yield insignificant results in both periods, suggesting that endogeneity problem may not occur under these specifications.

From the discussion above, the IV models, either by using policy instruments or smoking as a bad habit variable, confirm that there is a decline in the return to education between the year 2000 and 2014, as is in line with the OLS models.

Sample Selection Bias

There are two issues with regard to the sample selection bias. Firstly, the bias arose from non-random sampling for employment. Secondly, the bias is for female workers, as explained in Section 3.2.2. This part will focus on all individuals (selection on labour force participation). Meanwhile, Two-Step Heckman models for gender are provided in Appendix X.

To conduct the analysis, the present study follows Dumauli (2015) to define the instruments. Dumauli (2015) uses household assets, regional unemployment rate, marital status and household size. However, the OLS models in the present study already use marital status as a control variable and regional employment rate is also already represented as the regional control variable (province dummies). Household assets data are available in the survey but using this instrument would reduce the number of observations substantially. As such, this study considers using only household size, defined by the number of household members, to retain the number of observations. Household size can influence the decision (particularly for females) to join the labour market since these variables can influence females (likewise, males) in terms of housework load and the time that has to be allocated for their families.

Doan *et al.* (2016) further that household size may affect wage and employment participation probabilities through changing the opportunity cost of being in the waged labour force. However, an employer is unlikely to pay a different wage rate depending on one's household size or non-labour income. In the agriculture sector, household size may affect waged employment participation due to low productivity. Besides, limited arable land in the agricultural sector has resulted in labour surplus in the sector if households have more members. Therefore, household size relates to labour surplus and affects waged employment participation. The other alternative instrument is the number of dependent children in the household, some households reported there had children, some of them did not reported, this may imply zero children or missing data.

Ideally, the instrument is decided by pre-treatment measures. The key assumptions for a pre-treatment variable to be a valid instrument are: 1) the instrument variable is associated with the treatment; 2) the instrument variable is not associated with unmeasured confounders after conditioning on measured confounders (i.e., after controlling for measured confounders by regression or matching); 3) the affects the outcome only through the treatment (i.e., there is no direct effect of the on outcome; this assumption is known as "exclusion restriction") (Ertefaie, et al., 2017).

The present study uses household size as the instruments to retain the number of observations, follows previous studies such as: Dumauli (2015) and Doan *et al.* (2016), in particular Dumauli also study Indonesia case. In additions, agricultural sector in Indonesia is still relatively large, although the trend has decreased, around 23 per cent workers are in agriculture, forestry and fisheries (SAKERNAS, 2017). The other alternative instruments and pre-treatment measures can be elaborate in the future research.

Table 3.32 shows the summary statistics of households as an instrument. The main advantage of the instrument is that it can retain the number of observations; Two-Step Heckman models have 6,286 and 8,115 uncensored observations in 2000 and 2014 models, respectively. The mean value of household size decreases significantly from 6 members (on average) in 2000 to 4 members in 2014. This could imply that the more educated people tend to have fewer children and smaller household size.

Table 3.32: Summary Statistics of Household Size

Variable	Obs	Year	Mean	SD	Min	Max
Uncensored observations	6386	2000	6.02	3.00	1	23
Number of obs.	13307	2000	6.04	2.87	1	37
Uncensored observations	8115	2014	4.41	3.32	1	27
Number of obs.	18794	2014	4.50	3.22	1	27

Source: The author's calculation.

Tables 3.33 and Table 3.34 show the estimation result of the Two-Step Heckman for non-random sample for employment (all individuals). The coefficient of household size that is included in the Probit model has the expected sign (a negative sign) and is statistically significant for all individuals in 2000 and 2014; the negative sign implies that there is a negative relationship between the household size and the decision to join the labour market, as well as the trade-off between time that has to be allocated for their families and the time for working or joining the labour market. The coefficient of Inverse Mills Ratio (IMR) is negative and significant at 10 per cent for both years. A significant coefficient for IMR signals the presence of the sample selectivity, as explained in Section 3.3.2. Thus, selection problem exists in our model for all individuals according to this estimation.

In terms of the years of schooling coefficient, the results of Two-Step Heckman show a slightly lower value than the OLS. The returns to years of school (additional years of schooling) were 12.4 per cent and 9.9 per cent based on OLS model for 2000 and 2014, respectively. Based on Two-Step Heckman, the return in the year 2000 was 11.8 per cent, which declined to 9.8 per cent in 2014 (around 17 per cent decrease). This result also confirms that there was a decline in return to education between 2000 and 2014, as is in line with the OLS results, despite the sample selectivity problem in the models. In addition, the likelihood-ratio test (the LR test) is used to perform a test against the nested alternate model. The p-value of LR test is 0, thus, the null hypothesis is rejected on 5 per cent and 10 per cent levels of significance, implying that the Heckman selection equation with these data is useful and better than the standard OLS regression[27].

Table 3.33: Heckman Estimation, All Individuals: 2000

	2000								
	OLS			First Step: Probit			Second Step: OLS		
	Coef.	SE	P>t	Coef.	SE	P>z	Coef.	SE	P>t
HH Size				-0.022	0.008	***			
Years of Schooling	0.124	0.005	***	0.111	0.009	***	0.118	0.004	***
Experience	0.033	0.005	***	0.050	0.010	***	0.032	0.005	***
Experience squared	-0.001	0.000	***	-0.001	0.000	***	-0.001	0.000	***
Sex (1=female)	-0.328	0.025	***	-0.127	0.050	***	-0.328	0.024	***
Married and cohabitate	0.102	0.033	***	-0.038	0.075		0.100	0.033	***
Other (Separated, divorced and widowed)	-0.051	0.065		-0.016	0.128		-0.050	0.064	
Religion1: Islam	-0.017	0.107		-0.002	0.186		-0.012	0.106	
Religion2: Christian/Protestant	0.081	0.121		0.152	0.220		0.083	0.121	
Religion3: Catholic	0.047	0.128		-0.316	0.237		0.061	0.128	
Religion5: Buddhist	-0.184	0.199		0.778	0.359	**	-0.205	0.198	
Ethnicity1: Jawa	-0.043	0.047		0.022	0.093		-0.044	0.047	
Ethnicity2: Sunda	-0.064	0.057		0.221	0.123	*	-0.064	0.057	
Ethnicity3: Batak	0.058	0.094		-0.253	0.186		0.063	0.094	
Ethnicity4: Betawi	0.005	0.067		0.010	0.155		0.010	0.066	
Ethnicity5: Minang	0.052	0.093		-0.078	0.192		0.053	0.093	
Ethnicity6: Tiong Hoa	0.109	0.157		-1.526	0.274	***	0.151	0.158	

[27] The Heckman models for gender and private sector analysis is provided in Appendix X for further information. However, Two-Step Heckman model results are not always valid for gender analysis since household size (an instrument variable) is insignificant, particularly for the model for males in the year of 2000 and females in 2014. Moreover, considering the different characteristics of the public and private sectors, Heckman model cannot be applied on the sectoral model.

Status: fulltime (30 hours a week or more)	-0.646	0.032	***	0.435	0.055	***	-0.650	0.032	***
Tenure	0.022	0.005	***	0.073	0.009	***	0.020	0.005	***
Tenure squared	0.000	0.000	*	-0.002	0.000	***	0.000	0.000	
Sector: private	-0.311	0.038	***	3.682	0.063	***	-0.380	0.059	***
Industry2: mining and quarrying	0.360	0.121	*	0.277	0.294		0.355	0.120	***
Industry3: manufacturing	0.118	0.040	*	0.239	0.083	***	0.119	0.040	***
Industry4: electricity, gas and water	0.126	0.143		1.236	0.310	***	0.109	0.143	
Industry5: construction	0.337	0.049	***	0.207	0.110	*	0.329	0.049	***
Industry6: wholesale, retail, restaurants and hotels	0.070	0.045		-0.057	0.078		0.076	0.045	*
Industry7: transportation, storage, and communications	0.167	0.057	***	0.391	0.104	***	0.166	0.057	***
Industry8: Finance, insurance, real estate and business services	0.434	0.094	***	1.582	0.342	***	0.405	0.094	***
Industry9: Social services	0.114	0.038	***	1.259	0.069	***	0.092	0.040	**
Firm size2: 5-19 people	0.183	0.027	***	0.989	0.054	***	0.167	0.029	***
Firm size3: 20-99 people	0.249	0.032	***	1.441	0.080	***	0.222	0.034	***
Firm size4: >= 100 people	0.339	0.043	***	1.500	0.140	***	0.320	0.044	***
province1: Aceh	(omitted)			(omitted)			(omitted)		
province2: Sumatera Utara	-0.542	0.438		-0.698	1.616		-0.533	0.437	
province3: Sumatera Barat	-0.652	0.445		-0.676	1.622		-0.651	0.444	
province4: Riau	0.135	0.457		-0.138	1.645		0.148	0.456	
province5: Sumatera Selatan	-0.705	0.438		-1.009	1.613		-0.688	0.438	
province6: Bengkulu	(omitted)			(omitted)			(omitted)		
province7: Lampung	-0.752	0.441	*	-0.875	1.616		-0.747	0.440	*
province8: DKI Jakarta	-0.359	0.437		-0.968	1.614		-0.355	0.436	
province9: Jawa Barat	-0.483	0.437		-0.820	1.613		-0.481	0.436	
province10: Jawa Tengah	-0.738	0.437	*	-0.779	1.613		-0.735	0.436	*
province11: D I Yogyakarta	-0.882	0.439	**	-0.788	1.614		-0.877	0.438	**
province12: Jawa Timur	-0.722	0.437	*	-0.807	1.612		-0.720	0.436	*
province13: Bali	-0.765	0.446	*	-0.510	1.620		-0.758	0.445	*
province15: Kalimantan Tengah	-0.902	0.438	**	-0.516	1.612		-0.907	0.437	**
province16: Kalimantan Selatan	-0.166	0.545		1.287	1.993		-0.175	0.543	

	OLS Coef.	SE	P>t	Probit Coef.	SE	P>z	Second Step Coef.	SE	P>t
province17: Kalimantan Timur	-0.571	0.438		-0.249	1.612		-0.574	0.437	
province18: Sulawesi Selatan	-0.896	0.439	**	-0.637	1.612		-0.891	0.438	**
province19: Sulawesi Tenggara		(omitted)			(omitted)			(omitted)	
Urban	-0.004	0.026		0.019	0.052		-0.007	0.026	
IMR							-0.100	0.058	*
Constants	7.188	0.454	***	-3.503	1.624	**	7.337	0.460	***
Uncensored observations				6386			6386		
Number of observations				13307					
The R-squared statistic				0.37			0.37		
				LR chi2(48)	14684.07		F(48, 6337)	78.09	
				Prob > chi2	0.00		Prob > F	0.00	

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table 3.34: Heckman Estimation, All Individuals: 2014

	2014: All Individuals								
	OLS			First Step: Probit			Second Step: OLS		
	Coef.	**SE**	**P>t**	**Coef.**	**SE**	**P>z**	**Coef.**	**SE**	**P>t**
HH Size				-0.015	0.004	***			
Years of Schooling	0.099	0.004	***	0.012	0.005	**	0.098	0.004	***
Experience	0.020	0.004	***	0.017	0.005	***	0.019	0.004	***
Experience squared	0.000	0.000	***	0.000	0.000	**	0.000	0.000	***
Sex (1=female)	-0.266	0.022	***	-0.507	0.030	***	-0.238	0.026	***
Married and cohabitate	0.057	0.032	*	-0.014	0.046		0.057	0.032	*
Other (Separated, divorced and widowed)	0.107	0.059	*	0.304	0.078	***	0.089	0.060	
Religion1: Islam	0.058	0.091		-0.096	0.120		0.059	0.091	
Religion2: Christian/Protestant	0.252	0.123	**	-0.206	0.175		0.261	0.123	**
Religion3: Catholic	0.213	0.111	*	-0.269	0.145	*	0.226	0.111	**
Religion5: Buddhist	0.103	0.301		0.122	0.459		0.090	0.301	
Ethnicity1: Jawa	0.132	0.036	***	0.053	0.050		0.129	0.036	***
Ethnicity2: Sunda	0.030	0.046		0.114	0.068	*	0.025	0.046	
Ethnicity3: Batak	0.020	0.076		-0.237	0.097	**	0.033	0.076	
Ethnicity4: Betawi	-0.092	0.164		0.061	0.240		-0.102	0.164	
Ethnicity5: Minang	0.155	0.072	**	-0.118	0.091		0.161	0.072	**

Ethnicity6: Tiong Hoa	0.367	0.166	**	-0.467	0.224	**	0.395	0.167	**
Status: fulltime (30 hours a week or more)	-0.461	0.029	***	0.347	0.031	***	-0.487	0.031	***
Tenure	0.019	0.004	***	-0.027	0.005	***	0.021	0.004	***
Tenure squared	0.000	0.000		0.000	0.000	***	0.000	0.000	
Sector: private	-0.290	0.034	***	0.816	0.015	***	-0.334	0.040	***
Industry2: mining and quarrying	0.370	0.087	***	0.732	0.139	***	0.295	0.093	***
Industry3: manufacturing	0.039	0.045		0.959	0.053	***	-0.045	0.060	
Industry4: electricity, gas and water	0.135	0.112		0.927	0.195	***	0.053	0.118	
Industry5: construction	0.302	0.062	***	0.779	0.065	***	0.226	0.072	***
Industry6: wholesale, retail, restaurants and hotels	-0.040	0.046		0.440	0.044	***	-0.088	0.051	*
Industry7: transportation, storage, and communications	0.099	0.072		0.793	0.091	***	0.023	0.080	
Industry8: Finance, insurance, real estate and business services	0.222	0.056	***	1.185	0.093	***	0.131	0.071	*
Industry9: Social services	-0.140	0.044	***	1.583	0.045	***	-0.260	0.072	***
Firm size2: 5-19 people	0.068	0.029	**	1.131	0.033	***	-0.029	0.054	
Firm size3: 20-99 people	0.244	0.031	***	1.697	0.050	***	0.123	0.065	*
Firm size4: >= 100 people	0.514	0.035	***	1.823	0.064	***	0.390	0.068	***
Province 1: Aceh (11)	1.319	0.916		0.237	3.488		1.321	0.916	
Province 2: Sumatera Utara (12)	0.387	0.071	***	0.138	0.089		0.379	0.072	***
Province 3: Sumatera Barat (13)	0.358	0.090	***	0.406	0.113	***	0.333	0.091	***
Province 4: Riau (14)	0.721	0.130	***	0.686	0.173	***	0.673	0.132	***
Province 5: Jambi (15)	0.546	0.268	**	-0.041	0.311		0.539	0.268	**
Province 6: Sumatera Selatan (16)	0.349	0.068	***	0.239	0.078	***	0.340	0.068	***
Province 7: Lampung (18)	0.148	0.077	*	0.103	0.087		0.144	0.077	*
Province 8: Kepulauan Bangka Belitung (19)	0.564	0.110	***	0.123	0.159		0.556	0.110	***

Province 9: Kepulauan Riau (21)	0.886	0.184	***	0.375	0.333		0.869	0.184	***
Province 10: DKI Jakarta (31)	0.667	0.058	***	0.210	0.082	***	0.657	0.058	***
Province 11: Jawa Barat (32)	0.408	0.061	***	0.081	0.081		0.403	0.061	***
Province 12: Jawa Tengah (33)	0.088	0.061		0.173	0.078	**	0.077	0.061	
Province 13: D I Yogyakarta (34)	0.035	0.069		0.059	0.093		0.032	0.069	
Province 14: Jawa Timur (35)	0.115	0.058	**	-0.082	0.072		0.120	0.058	**
Province 15: Banten (36)	0.612	0.068	***	0.082	0.101		0.609	0.068	***
Province 16: Bali (51)	0.593	0.094	***	0.233	0.127	*	0.576	0.095	***
Province 18: Kalimantan Barat (61)	1.344	0.460	***	(omitted)			(omitted)		
Province 19: Kalimantan Tengah (62)	1.264	0.242	***	-0.076	0.344		1.250	0.242	***
Province 20: Kalimantan Selatan (63)	0.400	0.084	***	0.323	0.101	***	0.384	0.084	***
Province 21: Kalimantan Timur (64)	0.913	0.143	***	0.462	0.240	**	0.882	0.143	***
Province 22: Sulawesi Selatan (73)	0.321	0.065	***	-0.003	0.081		0.323	0.065	***
Province 23: Sulawesi Tenggara (74)	0.205	0.266		0.209	0.316		0.183	0.266	
Province 24: Papua Barat (91)	(omitted)			(omitted)			(omitted)		
Urban	0.139	0.024	***	0.279	0.032	***	0.123	0.026	***
Mills							-0.188	0.089	**
Constants	7.862	0.163	***	-3.652	0.165	***	8.247	0.244	***
Uncensored observations			8119					8115	
Number of observations					187934				
The R-squared statistic			0.303					0.30	
				LR chi2(53)	15150.85		F(53, 8060)	66.2	
				Prob > chi2	0		Prob > F	0	

Source: The author’s calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

3.5 Conclusion

Throughout the literature on education and wages in Indonesia, most studies focus on the relationship between wages, education levels or years of schooling, experience, gender, marital status and areas or residence. The Mincer wage equation model in the present study is relatively more systematic and comprehensive (with many control variables included) than the previous studies in Indonesia.

This chapter has focused on three research questions, as addressed in section 3.1. The first research question is: “What are the estimated return to education in the year 2000 and 2014?” Some alternative Mincer wage equations are estimated in this study, and the result is consistent among those estimations. Overall, the return to education varies and increases in line with education level, as predicted by the human capital theory. More specifically, the return to junior high school in the year 2000 (relative to primary school and below level) was 28.5 per cent, senior high school was 63.7 per cent and university was 119.3 per cent. In 2014, the return to junior high school was 22.1 per cent and to senior high school and university were 48 per cent and 95.8 per cent, respectively. In terms of years of schooling, the result shows that one additional year of schooling increased wages by 12.4 per cent in 2000 and by 9.9 per cent in 2014.

Regarding the effects of gender on the return to education, there is a gender disparity in wages, with the return to education for females being higher than males, particularly for the senior high school and university graduates. The results also show the same pattern; the return to education increases in line with education level for both males and females, confirming the findings of Psacharopoulos (1994), Harmon *et al.* (2000), and Dumauli (2015). Some possible explanations are the self-selection, a more limited supply of skilled female workers, and the different technological requirements in female-dominated and male-dominated jobs (Ren and Miller, 2012).

In terms of sectors, the return to education is generally higher in the public sector than in the private sector for both periods, similar to Filmer and Lindauer (2001). But it appears that public and private sectors in developing countries are non-competing, as stated in intersectoral linkages in the labour market (Fields, 2010). Some possible reasons for the premium in the public sector are: the increase of wage in the public sector is higher and faster (The World Bank, 2000), around 60 per cent of the sample chose to settle in urban areas which tend to provide higher wages than the rural employment, there is a lack of competition in the public sector (Psacharopolous,1979), and the private sector is more efficient than the public sector (Rao, 2015), especially in productive and allocative efficiency[28], thus wage rigidity may occur more in the public sector.

The last research question examines the change in the return to education between those periods. The results show that the return to education tends to decline for most education levels between 2000 and 2014. This agrees with Gropello and Sakellariou (2010), even though the result is slightly different to Purnastuti *et al.* (2013) who assert that the profitability of university qualifications increases between the years 1993 and 2007. Moreover, the expansion in education significantly affects the return of lower levels of education (junior high school relative to primary school). This could be attributed to the 9-year compulsory education program's effect which causes a surge in the net enrolment ratio (NER); thus, more individuals have at least junior high school qualifications. On the other hand, more companies prefer workers with higher qualifications, as Allen (2016) asserts. In contrast, Dumauli (2015) argues that the decline in return to education in Indonesia is probably because high skilled jobs are not in demand in Indonesia, and the low quality of the education system affects the quality of the graduates. As such, in terms of sectors, the private sector experienced a significant decline in the return to education between the year 2000 and 2014.

Anticipating that endogeneity problems and selection bias may occur in the estimations, the present study performs the robustness tests. The IV results show that endogeneity problems do not occur, while the Heckman model confirms that selection bias could occur in the OLS models. Despite the instruments' limitation, the result of IV and Heckman model indicate that the return to education increases in line with education level. Those models also confirm that there is a decline in the return to education in Indonesia between the year 2000 and 2014. More specifically, the OLS returns are slightly higher than the Heckman and IV results. Since the patterns of return to education are similar among the OLS, the IV and the Heckman model; and the limitation of the IV and the Heckman model, the present study prefers the OLS model.

[28] Productive efficiency refers to the maximisation of outputs over inputs or doing the most work with the fewest resources. Allocative efficiency refers to the match between the demand for services and their supply or allocating resources to the right place to do the right job (UNDP Global Centre for Public Service Excellence, 2015).

In terms of other control variables, ethnicity caused a different effect; generally insignificant in the year 2000, but some ethnicities, such as Jawa, turned significant in 2014, thus suggesting that being a majority ethnicity has an advantage on wages. Similar to previous studies, the present study evidences a positive relationship between firm size and wages. Furthermore, working in non-agricultural industries still offered higher wages (relative to agriculture) in the year 2000. However, as labour-intensive industry experienced a significant decline (Aswicahyono *et al.*, 2011), wages from labour intensive industries (such as: manufacturing; wholesale, retail, restaurant and hotels) became insignificantly different from wages in the agriculture sector in 2014. In terms of provinces and areas, living on the Jawa island and urban areas could provide higher wages, thus still being an appeal for a migration into those areas.

This research has some implications, especially for individuals who choose the public or private sector for his/her future career, higher education offers the higher wages based on the model, despite the decline in the return to education overtime. For the government, investing in higher education and improving the quality of the educational institutions are necessary. Nonetheless, as Irandoust (2013) asserts, despite the increase in years of schooling and greater overall participation in higher education, graduates are found to be unprepared for the job market. Nowadays, a university qualification does not necessarily guarantee that individuals will fit the needs of the industries or workplaces.

Moreover, as the public sector provides higher hourly wages than the private sector, this could create a crowding-out in the economy. Behar and Mok (2013) emphasize that crowding-out can occur through many channels, such as the labour market where higher wages, more job security, or a higher probability of finding a public-sector job can make an individual more likely to seek or wait for a public-sector employment rather than searching for or accepting a job in the private sector. Also, individuals seek qualifications appropriate for entering the public sector through the education market rather than seeking the skills needed for productive employment in the private sector. Slightly different to this few, according to intersectoral linkages in the labour market, both public and private sector is a noncompeting group in which individuals belong to one labour market segment or another, and they cannot or will not switch from one to another.

This research also has some limitations that can be improved in further studies. The first one is related to the IV instruments. As policy instruments have some limitations, the concern is to retain the number of observations of the OLS. Other alternative instruments can be explored for better estimations in further research. The model also could have included casual workers, since their wage data are available in IFLS5 (2014), hence the analysis could have given a more comprehensive picture of the return to education in Indonesia. And finally, one finding of this chapter is that the excess supply of college graduates turns to reduce the wages for them, which could relate to education mismatch in the labour market. Such issues will be analysed in the next chapter.

Chapter 4 Education Mismatch in Indonesia

4.1 Introduction

Chapter 3 of the present study finds that there is an expansion in education, especially at high school and university levels, as The Gross Enrolment Ratio (GER) of both levels increases significantly, as indicated in Figure 2.2. In contrast, there is a decline in the relative wages of university and senior high school graduates (relative to primary school and lower qualifications). Freeman (1976) finds that overeducation occurs when the supply of university graduates increases more rapidly than its demand. Meanwhile, the human capital theory suggests that wages will always be equal to the individual worker's marginal product. However, firms need to adjust production process in order to fully utilise their human capital in the short run. Similarly, workers need time to find a more appropriate match. Thus, it is possible to suggest that education mismatch has occurred in Indonesia, and this will be discussed further on in this part.

Education mismatch arises when the educational qualifications of the workers are different from the qualifications required by their jobs, therefore the phenomenon can either be an undereducation or an overeducation. In Indonesia, Nazara and Safuan (2005) estimate this mismatch using Sakernas data and find that undereducation decreased from 16.8 per cent in 1999 to 9.13 per cent in 2002, while overeducation increased from 26.7 per cent to 34.7 per cent in the same period. Using the same data source for later periods, ILO (2017) also finds that there was a decrease in overeducation from 27 per cent in 2006 to 19.2 per cent in 2016; for undereducation, the trend increased from 10 per cent to 17 per cent for the same period. Comparing the latest periods, both overeducation and undereducation proportion in Indonesia are significantly higher than both proportions in OECD; the average of undereducation for OECD countries is around 19.1 per cent, while the average of overeducation is 14.9 per cent in 2015.

Allen (2016) argues that the education system is not providing enough graduates in some sectors in Indonesia, while in other sectors those who are graduating do not have the right skills. Consequently, a large number of individuals with post-secondary qualifications work in low-skill occupations, which suggests that despite holding degrees, they lack the right skills. Allen also finds that 52 per cent of the employed population was underqualified for their positions in 2015. McKinsey Global Institute (2012) estimates that the demand for semi-skilled and skilled workers may rise to 113 million by 2030, which is likely to see skills shortages and skills mismatches worsen throughout the economy.

Hence, this chapter aims: to investigate the education mismatch and to update the existing literature on education mismatches, particularly in the waged sectors in Indonesia during the 2000 and 2014 periods; to analyse variables that affect the mismatch in the waged sectors, and; to explore the changes on the mismatch determinants between these periods. The present research seeks to answer the following research questions: (4.1) Does education mismatch (both undereducation and overeducation) exist in the waged sectors in Indonesia? (4.2) What are the estimates of education mismatch in 2000 and 2014? How does the aggregate trend of education mismatch change between these periods? Are there any distinctions among genders and sectors? (4.3) What are the variables which determine undereducation and/or overeducation? And are there any distinctions among genders and sectors?

Similar to Chapter 3, this chapter uses IFLS3 of 2000 and IFLS5 of 2014 to provide answers to the aforementioned research questions. Furthermore, since the rapid expansion and reform in education occurred in mid 2000s, thus the 2000 data represent the situation before the reform, and the 2014 data (the latest IFLS data available) represent situation after the reform. To estimate the education mismatch, this research employs objective measures (calculating mode and standard deviation) or realised method (RM)[29]. In measuring the mismatch, the mode[30] years of education could be a better instrument compared to the mean level of education because mode is commonly used for categorical data like mismatch data. Furthermore, mode is a better measure than the mean when the data are skewed or not normally distributed (Lund and Lund, 2018). In fact, data with normal distribution will have equal mean, mode and median.

Besides the data source, time period, and method (mode or mean), another aspect that can offer different results of mismatch estimations is occupational segregation[31]. IFLS data

[29] Commonly, most of previous studies uses mean. Mode is in terms of year of education.

[30] Mode refers to the most frequent data which appear in the dataset.

[31] According to the International Standard Classification of Occupations (ISCO-08), occupation groups are divided by major groups (1 digit), sub major groups (2 digits), minor groups (3 digits), unit groups (4 digits occupation categories). When a mismatch is measured by the mean of occupation, different levels of occupation category may have different mean; as a result, the mismatch estimations may differ.

provide general categories (1 digit)[32] and sub-categories (2 digit); using either 1 or 2 digit generates several issues such as a wide range of years of schooling within the occupation category for 1 digit and a low number of observations for the 2 digits. As such, the present study rearranges the occupation group (the hybrid category) in order to reduce heterogeneity and to avoid low number of observations, by keeping sub-categories which have relatively large observations, and merges the remaining (particularly sub-categories with sample size of less than 30). Compared to previous studies such as Nazara and Safuan (2005) and ILO (2017), this occupation group (hybrid) is relatively more detailed. Likewise, Foxton (2016) asserts that using an appropriate digit level of occupation category can provide a good balance between a strong sample size and reducing the level of heterogeneity in roles within occupational grouping.

The present research develops a mismatch determination model that consists of personal characteristic variables, household characteristics, work-related and firm size variables and residence or area dummy variables. Multinomial Logit Model (MNL) is also applied to estimate the mismatch model. The main advantage of using MNL is its capability to analyse more than two possible discrete outcomes, *i.e.* both overeducation and undereducation, and the simpler logit model is also often preferable to the more complex probit model (Dow and Endersby, 2004). The waged sector is chosen for this study considering that the characteristics of public and private sectors are relatively similar; they have regular wages and the data are available. Moreover, the waged sectors, especially the public sector, are very attractive for job seekers, for example; the ratio of job opportunities in the public sector to the number of applicants is around 1:200 (Sindo, 2013).

The present study also extends the analysis based on sector and gender to examine the different effects each variable has on gender and sector, since different characteristics of gender and sector may have a substantial effect on matches. For instance, ILO (2017) reveals that male workers in Indonesia tend to have lower education levels than the female workers. However, males receive relatively higher wages than females. Public sectors also have a relatively higher return to education than the private sectors (see findings in Chapter 3). In addition, there is a change in regulation for example, the Government

[32] 1-digit IFLS categories are similar to the ISCO-08 major occupation group from ILO, which were commonly used by previous studies such as: Allen (2016), Chua and Chun (2016), Lee *et al.* (2016) and Flisi et al. (2017).

Regulation 30/2015 on list of basis salaries of civil servants, the basic salaries is based on entry ranks, which are mainly determined by education level, and increases in rank are largely driven by seniority, civil servants in Indonesia are divided into four ranks, from I (the lowest) to IV (the highest). Rank I through III are divided into four grades (a, b, c, and d), and Rank IV has five grades (a, b, c, d, and e), making a total of 17 grades from I (a) to IV (e). Individual civil servants' ranks are based on their educational qualifications and seniority. Ranks III and IV require and university degree, while Rank I only need an elementary and secondary education and Rank II mostly occupied by who has finished a senior high school (Tjiptoherijanto, 2018). The regulation is revised every year to adjust the civil servants wage with the inflation that applies to both of new entrants and the remaining civil servants. This may affect the increase of overeducation incidences in Indonesia since workers prefer to have at least university qualifications and they can entry with rank III with higher wage standard. In the other side, the public sector also prefers to hire workers with at least senior high school qualifications. While, for unskilled jobs such as cleaning, the public sector prefers outsourcing or using private firms' services rather than hiring directly. The other regulations that may affect education mismatch such as: the minimum requirement for lecturer is PhD in top five universities. Such regulations can be addressed in the future research.

The present study will also provide results of a sensitivity test by providing some alternative methods: (1) using mean, as previous studies have commonly practised; (2) using multinomial probit (MNP); and (3) extending the sample by adding casual workers for 2014, to see if the model still applies and is consistent with the main model (the waged sector). As these casual workers are added, analysing both the waged sector and casual workers can provide a more comprehensive picture on Indonesia's labour market, since around 60 – 70 per cent of the workforce is estimated to engage in informal employment, which include casual workers (Statistics Indonesia, 2010). Unfortunately, the model with casual and waged workers cannot be applied in the year 2000 because the casual workers' data are not available in IFLS3.

Therefore, Chapter 4 is structured into: Section 4.2, providing a summary of the Indonesia's context; Section 4.3, reviewing various theories and the determinants of education mismatch, including the change of mismatch; Section 4.4 explaining the data and methods of the study; Section 4.5, discussing the estimation result and test of sensitivity results and finally; Section 4.6, presenting the conclusion of this chapter.

4.2 Literature Review

4.2.1 Mismatch Definitions and Measurements

Mismatch Definitions

Freeman (1976) introduced the notion of education mismatch, particularly overeducation, in which education is measured objectively by comparing a worker's level of education attainment with what is required by the worker's job. Consequently, education mismatch arises when the educational qualifications of the workers, individually or in the aggregate, are different from the qualifications required by or specified for their jobs (Sattinger, 2012), and as such it can be an over/undereducation. Overeducation refers to the phenomenon where workers have more education than their job requires (Silles and Dolton, 2002), while undereducation occurs when workers have lower education than the requirement. In other words, a match occurs when workers have exactly the education the job requires. Here, education is defined as the highest level or years of education achieved by individuals.

Conceptually, many studies have also explored different definitions of education mismatch, such as Rumberger (1981) who defines the mismatch in three ways: firstly, as a decline in the economic position of educated individuals relative to historically higher levels; secondly, as an under-fulfilled expectation of the educated with respect to their occupational attainments; and thirdly, as the possession of greater educational skills than their jobs require. Furthermore, Oliveira *et al.,* (2000) argue that undereducation is the outcome of a process in which market-acquired capital substitutes for insufficient school-supplied qualifications, where overeducation is associated with excess schooling but short tenure and job experience. In a similar vein, Gosling and Zhu (2010) separate overeducation definitions based on micro and macro levels. In micro level, overeducation is defined similar to Silles and Dolton's (2002) definition. Meanwhile, in macro level, there are some characteristics of overeducation, such as a labour market which has "too many" graduates, a credentialism tendency and thus represents a disequilibrium.

As a result, a match is the most efficient condition for the economy, education mismatch, overeducation in particular, is potentially costly to the economy, the firms and the individuals. At the macroeconomic level, national welfare is potentially lower than what would be the case if the skill of all under/overeducated workers were fully utilised within the economy. It is also (probably) related to tax revenues being wasted on equipping individuals with non-productive education. At the firm level, overeducation could be associated with lower productivity, as Tsang (1987) finds that overeducated workers have a negative effect on output. At the individual level, overeducated workers, by virtue of the fact that a proportion of their educational investment is unproductive, are likely to earn a lower return on their investment relative to similarly educated individuals whose jobs match their education (Ortiz, 2010).

According to previous empirical studies, overeducation measurement could also belong to the upper tail of the education distribution based on statistical definition. Similarly, undereducation refers to the lower tail of the distribution (Rumberger, 1981). Hartog (1997) defines overeducation as departing from more than one standard deviation from the mean, resulting in finding similar proportions of over and under educated workers - around 15 per cent of the population, if education is measured in years and the distribution of education per occupation is normally distributed. Kampelmann and Rycx (2012) find an increase of one year in the incidence of undereducation among young workers is found to decrease productivity on average by 3.5 per cent one year later.

In additions, another strand of literature on overeducation finds that there is a negative relationship between overeducation and job satisfaction (Battu *et al.*, 1999; Chevalier, 2003). Chevalier (2000) proposes an alternative measure of mismatch based on occupation and job satisfaction, more precisely, whether the graduate is satisfied with the match between her education and her occupation. Graduates in a sub-graduate occupation who are satisfied are defined as apparently over-educated, whereas those who are dissatisfied are called genuinely over-educated. One advantage of using this definition of mismatch is that it refers to qualifications, not only education, and does not require an assessment of the educational level which requires doing the same job. Yet, it can be argued that this definition does not measure overeducation accurately, because the dissatisfaction between qualification and occupation could be due to undereducation; the dissatisfaction could reflect that despite being in a graduate job, this occupation is not related to the academic subject studied at university; and/or the job may require most of the skills that were learnt at university but also some more from a different field. To answer these criticisms, the definition of overeducation (based on job satisfaction) should be combined with the Professional Job Analysis (JA) method. Similarly, the negative correlation between overeducation and job satisfaction is also confirmed by Clark (2014).

In a similar respect, Chevalier (2000) distinguishes overeducation into two parts: apparent overeducation (*i.e.* overeducated only) and genuine overeducation (*i.e.* both overeducated and mismatched skill-wise). Moreover, a mismatch can either be vertical or horizontal or both. A vertical mismatch occurs when the level of the employee's qualification is not the one required by the job, for instance, a graduate employee who works in a job that is typically considered a non-graduate job, in which case the graduate is over-educated. Meanwhile, a horizontal mismatch occurs when the level of the employee's qualification is at the correct level for the job, but the type of the qualification is not, for example an individual with a degree in engineering working in a job that requires no engineering knowledge at all. A horizontal and vertical mismatch occurs when an individual may have a qualification that is both at the wrong level and of the wrong type for the job they are hired to do.

It is also worth noting that education and skill are not synonymous[33]. As Flisi *et al.*, (2017) distinguish, education refers to an individual's qualifications at a given point in time, which are bound by differences across countries and cohorts for the same level attained. By contrast, skills are acquired and lost over an individual's entire lifespan, thereby providing a more concise and updated measure of competencies[34]. McGuinness *et al.* (2017) further that overskilled is a situation where a worker believes that they possess more skills than their current job requires, while underskilled occurs if the worker believes that their current skills do not meet the demands of the job. Mavromaras *et al.* (2009) argue that overskilled could be a more accurate measure of mismatch amongst existing workers than overeducation on the grounds that overeducation assumes that; (a) job entry requirements accurately reflect job skill content, and (b) a worker's qualifications adequately reflect their total work-related human capital. Thus, the overeducation approach ignores the fact that job entry requirements may be weakly

[33] Flisi *et al.* (2014) report that around 30 per cent of EU employment is overeducated (but not overskilled), 17 per cent of them are overskilled (but not overeducated), and around 15 per cent are simultaneously mismatched; both overeducated and overskilled.

[34] In terms of skill, World Bank (2010) has published an Indonesia skills report which reviews the main characteristics of demand skills, documents the existence of a possible skill mismatch between employer demands and available supply and the contribution of the education and training sector to the mismatch. The analysis is based on employer and employee skill survey, involving around 473 medium and large firms and 200 employees in manufacturing and services sector in five provinces. The report highlights; an evidence of serious gap in generic and technical skills which puts young workers particularly in critical situation; at the more macro level, the employer survey suggests that skills are not yet a binding constraint to business development though the skills-for-needs issue does nevertheless appear non-trivial and, subjective assessments of difficulties of matching needs with available skills provide further evidence that skills are becoming a problematic aspect in Indonesia's economy.

related to job content, and more reflective of qualification inflation and credentialism, while individual human capital will also consist of (non-formal and informal) skills acquired through labour market experience and training. Another reason why overskilled may be a more comprehensive measure of mismatch is that it compares all the skills and ability, irrespective of whether they are learned in the classroom or work environment, with the actual skill requirements of the worker's current job (McGuinness *et al.*, 2017).

Although some studies find overskilled more accurate and comprehensive, there are several drawbacks of the method, such as: overskilled and underskilled are measured through separate questions, unlike education mismatch where a single question can be used to identify both over and undereducation. Also, skill mismatch is commonly measured through direct assessment by human resource specialists, and such direct measures are rarely captured in datasets. The questions adopted to investigate overskilled/underskilled vary substantially across datasets, consequently making it difficult to compare the estimates. Overskilled questions also could not allow the researcher to identify the relative importance of underused skills deriving from labour market experience, training, innate ability or formal schooling. In addition, skill mismatch measurement also could pose a bias in the estimates, for instance, the respondents' job is related to their hobbies, as such when they formulate the response, their consideration of skills and abilities is totally unrelated to the workplace (McGuinness *et al.*, 2017).

Mismatch Measurement

There are some basic approaches to measure a mismatch. These choices are typically restricted by data availability. Nonetheless, there is a growing literature centred on assessing the levels of consistency and potential biases associated with the various approaches. The summary of mismatch measurement is shown in Table 4.1, and the explanation follows after the table.

Table 4.1: Mismatch Measurement

Type of mismatch	Measurements	Approach		Method
Education Mismatch	Objective Measures	Normative/Professional Job Analysis (JA)		Using information provided by professional job analysts: comparing job titles with actual education attainments
		Statistical/Realised Match (RM)		Comparing education attainments with the mean of education level (± standard deviation)
	Subjective Measures	Self-declared/ Self-reported/ Self-assessment	Direct measures (DSA)	Asking respondents directly whether they are over/undereducated/matched
			Indirect measures (ISA)	Asking respondents to give information on minimum job requirements and individual's acquired education
		Mixed/ Alternative methods (EMX)		Mixed between those methods

Adapted from: Filsi *et al.*, 2014.

(1) Objective measures

Overeducation can be assessed objectively: (1) the Job Analysis (JA) method; by using the information provided by professional job analysts/JA (such as in the Standard Occupational Classification System in the UK or the Dictionary of Occupational Titles in the US) to determine an individual’s required education on the basis of their job title and again comparing this with their actual level of education (Rumberger, 1987); (2) Statistical/Realised Match (RM); by calculating the mean education level for a range of occupations with an individual defined as being overeducated if their qualifications are more than one standard deviation above their occupation’s mean education level (Sicherman, 1991); while undereducation occurs if the actual education attainment is lower than the mean level of education within their occupation (known as Realised Matches/RM). Apart from the mean, median or mode of educational level in an occupation can be used as well (Kiker *et al.*, 1997).

These measures are also open to criticisms such as: (1) some occupations may contain a number of skill levels, so that in fact people with the same job titles may be doing very different jobs, for instance, the tasks undertaken by managers are likely to vary widely; and (2) rising education levels in the economy imply that employers will allocate workers differently. For example, Mason (1996) reports that managers are now employing university graduates in mid-clerical positions, the posts traditionally held for persons educated to O and A levels standard (predominantly high school graduates). Thus, the educational requirements of various occupations will evolve with changes in relative supply; a factor that is not always readily incorporated into occupational classification systems that tend to be relatively static in nature. Hartog (2000) adds that a carefully conducted job analysis method should not lead to any systematic bias. Yet, this requires a regular update of the classification scheme. Otherwise, a general upgrade of skill requirements due to skill-biased technological change might lead to overestimation of the incidence of overeducation.

In terms of empirical evidence, Kiker *et al.* (1997) study the determinants of overeducation and undereducation in Portugal using Personnel Records (Quadros de Pessoal) dataset in 1991, collected by the Portuguese Ministry of Labour. Years of schooling variable is used to calculate the mismatch. The study employs three alternative methods to measure overeducation and undereducation: (1) the Verdugo and Verdugo (1989) or VV model, where job requirement is defined as actual occupation attainments of workers within occupations disaggregated at a 3-digit level. Workers whose education attainments fall within plus or minus one standard deviation of the mean value within the occupation are considered to be adequately educated; (2) the Mode model, where educational attainments equal the modal (mode) value within each occupation, overeducation or undereducation equals to modal education level plus or minus the standard deviation; and (3) the measurement based on job analysts' opinion. The study finds that overeducation and undereducation exist in the Portuguese labour market. Comparing these methods, over and undereducation are highest under the third method, around 33.1 per cent and 37.5 per cent of the sample, respectively. Meanwhile, the second method (mode) results are only 25.5 per cent for overeducation and 17 per cent for undereducation. The lowest estimation is reported by VV method, only 9.4 per cent for overeducation and 5 per cent for undereducation. The study also reports another period (1985) of the same data and by using the same method finds that there is an increase of overeducation from around 18 – 26 per cent in 1985 to 25 – 33 per cent in 1991 and a similar increase for undereducation as well. Furthermore, overeducated workers are more likely to be the young members of the employed labour force, while undereducated workers are more likely to be the older members. The finding from Kiker *et al.* (1997)

seems to be more supportive of the role of technological change rather than of human capital in explaining overeducation/undereducation for the Portuguese economy; characterised by intensive efforts to promote economic growth, modernization of the industrial structure, and the upgrade of educational qualifications.

(2) Subjective measures

Overeducation can also be measured subjectively: (1) the Indirect Self-Assessment (ISA); by asking the respondents to give information on the minimum requirements of their jobs and then comparing this with the individual's acquired education - some studies employed this method such as Duncan and Hoffman (1981) and Hartog and Oosterbeek (1988); or (2) the Direct Self-Assessment (DSA); by simply asking the respondents whether or not they are overeducated - used by Halaby (1994). And the last method is mixed or alternative method (EMX); by combining two or more methods above (Chevalier and Lindley, 2009).

There are some criticisms against these measures: (1) overeducated workers may be less likely to respond to questionnaires due to higher levels of job apathy which may lead to underestimation of the incidence of overeducation; (2) workers in smaller and/or less structured organisations may lack sufficient benchmarks against which they can assess their job requirements, a factor which will again lead to measurement error; and (3) even where benchmarks are available, respondents may be applying differing criteria when assessing their job requirements, *i.e.* the actual level of education required to do specific tasks or the formal educational requirements necessary to get the job.

Hartog (2000) highlights the obvious bias in self-assessment for the measure that is based on the assessment of the required level to get the job. These measures provide an indication of the credentials gap (Livingstone, 1998); nevertheless, the concept of credentials and overeducation is not exactly the same. For instance, employers might increase hiring standards in response to cyclical or structural oversupply of educated workers, possibly causing the level required to get the job to deviate from the actual level required to do the job. Conversely, if the self-assessment measure is based on this last level, the bias is likely to be less severe. Nonetheless, there might still be a problem if individuals tended to inflate the status of their own position or if they adapted their answers to their personal ambitions and expectations. These biases might also be a problem for measures that are based on DSA. Moreover, these indicators might measure skill mismatches instead of education mismatches, particularly if they were based on questions regarding skill utilisation.

In addition, Groot and Brink (2000) study education mismatch in the US and EU, using meta-analysis. By analysing 25 studies, they obtain 50 estimates on the incidence of overeducation and 36 estimates on the incidence of undereducation. They find that the unweighted average of the overeducation incidence is 23.3 per cent (with a standard deviation of 9.9 percentage points), while the unweighted average incidence of undereducation is 14.4 per cent (with a standard deviation of 8.2 percentage points). The unweighted averages of the rates of return to the different educational components are: 5.6 per cent for years of education attained, 7.8 per cent for years of education required for the job, 3.0 per cent for years of overeducation and 21.5 per cent for years of undereducation. They also find that the different definitions lead to large differences in the incidence of overeducation. Overeducation rate in the EU also appears lower than the US; the average value of overeducation among studies for the United States is 26.3 per cent, compared to 21.5 per cent among the European studies. In terms of aggregate trend, they also point out that the trend of overeducation and undereducation declines between 1970s and 1990s. With the decline in the incidence of overeducation over time, the average rate of return to years of overeducation has declined as well. With a similar trend for undereducation, subsequently, the joint decline in overeducation and undereducation suggests that skill mismatches in the labour market have decreased since the 1970s.

McGuinness (2006) also documents many studies on mismatch for various education levels, particularly overeducation in the US, Canada, Hong Kong, the UK and seven European countries[35], using RM and years of schooling as education attainment variable. The study concludes that objective measures with RM (mean) is found to generate lower estimate; around 22 per cent, some 7 per cent points below the comparable subjective figure. For cross-country comparisons The Netherlands yields the lowest incidence level under both subjective and objective measures (Hartog and Oosterbeek, 1988; Groot and Brink, 2000), whilst studies of the US labour market generate the highest incidence level irrespective of the measurement approach adopted (Tsang *et al.*, 1991; McGoldrick and Robst, 1996).

[35] Germany, the Netherlands, Greece, Spain, Portugal, Ireland and North Ireland.

In a similar vein, Chua and Chun (2016) study the mismatch in Asian countries, using the World Bank survey data for Armenia, the People's Republic of China (PRC, Yunnan), Georgia, the Lao People's Democratic Republic (Lao PDR), Sri Lanka, and Viet Nam between 2012 and 2013. The sample per country comprise around 3,000 adults aged 15 to 64 who were located in urban areas. The method used to estimate the mismatch is ISA. The study finds that urban Viet Nam has the highest share of individuals working in jobs for which they are overeducated. Undereducated incidence is most prevalent in Armenia, where 29.1 per cent lacks years of schooling to meet the job requirements.

4.2.2 Mismatch Theories

There are at least nine theories relevant to education mismatch, as indicated in Table 4.2.

Table 4.2: Summary of the Main Features of Mismatch Theories

Theory	Authors	Persistence	Main characteristic
Human capital theory	Becker (1964)	Temporary	The mismatch can easily be solved via individuals or firm adjustment
Matching theory	Pissarides (2000)	Temporary	Both individuals and firms look for matches
Job competition model	Thurow (1975)	Persistent	Labour market allocation is based on hierarchy of workers and jobs' education level
Assignment theory	Sattinger (1993)	Temporary or persistent	Takes into account individuals' preferences on job/sector/wage maximisation
Career/job mobility theory	Sicherman and Galor (1990)	Temporary or persistent	Individuals are unable to properly signal their skills become overeducated
Search frictions	Gautier *et al.* (2010)	Temporary	Lack of information in the beginning of one's career
Signalling / screening	Spence (1973)	Persistent	Informational asymmetry exists in the labour market, education acts as an important signal
Preferences	Gottschalk and Hansen (2003)		The same productivity exists among workers, college workers voluntarily choose to work in the non-college sector in equilibrium
Technological change theory	Mendes De Oliveira *et al.* (2000)	Temporary or persistent	Rapid technological change increases the undereducated, both individuals and firms look for matches

Adapted from: Capsada-Munsech, 2017.

The job competition theory offers a demand side explanation, in contrast to the supply side approach of the human capital and career mobility theories. There are several alternative theories to explain the mismatch such as: the assignment models, search and frictions, signalling/screening, technological change theory, and preferences. However, this part only elaborates some of the key theories that have been used in most of education mismatch studies, *i.e.* the human capital, job competition theory, assignment, the career mobility (Linsley, 2005), signalling and screening, and technological change theory.

Human Capital Theory

The discussion under this section focuses on the human capital theory in relation to education mismatch or overeducation. As briefly explained in part 3.2.1, the human capital theory argues that education endows an individual with productivity-enhancing ability, and workers will always be paid by their marginal product (Becker, 1975). One of the propositions of the theory is that firms are willing to fully utilise the skills of their workforce by adapting their production processes in response to any changes in the relative supply of labour. Consequently, wages will always be equal to the individual worker's marginal product, which in turn will be determined by the level of human capital that they have accumulated through either formal education or on-the-job training. Also, overeducation or undereducation will not exist in equilibrium since there will be no under-utilisations of human capital in the labour market.

According to the theory, overeducation is associated with worker's under-utilisation and wage rates which are below the marginal product; overeducation would appear entirely inconsistent with this view of the labour market. However, as overeducation will become apparent, some economists have continued to argue that the human capital theory remains fully consistent despite the existence of overeducation. Overeducation phenomenon does not certainly overturn human capital theory as it is entirely plausible that workers will be overeducated in the short run (temporary), at the same time as firms adjust their production processes in order to fully utilise the individuals' human capital or alternatively for as long as it takes workers to find a more appropriate match through job search. Thus, the human capital theory can be rationalised by allowing for the existence of short-run disequilibria.

The explanation of overeducation is also provided by empirical framework, adopted for testing the human capital theory; this could also be entirely consistent with the neoclassical view. The standard approach was developed by Mincer (1974) and is based on wages regression centred on years of schooling which is replaced by overeducation variables; less formal measures of human capital such as on-the-job training, which

Becker (1975) argues is directly substitutable with schooling, are ignored in this case. Thus, this specification implies that individuals with more schooling may be compensating for a lack of work-related human capital, and the apparent lower wages of these 'overeducated' may be attributable to an omitted variables problem, such as a lack of controls for less formal measures of human capital accumulation. Furthermore, it may also be the case that overeducated workers are in some way less able relative to their adequately matched counterparts; hence, lower wages are merely a reflection of lower ability and/or productivity.

In addition, Freeman (1976) examines overeducation using US universities data and empirically finds that the long-run mismatch arose in the US in 1970's because the supply of university graduates increased more rapidly than the demand for it. As a result, many individuals with higher education levels were unable to find jobs that require something similar to their education attainment. Long-run mismatch also contributes to the short-run mismatch in the form of overeducation for many individuals. Nonetheless, even in the absence of any long-run mismatches, short-run mismatches would continue to exist due to the difficulty of finding a job without mismatches in a reasonable amount of time (Sattinger, 2012).

Kiker *et al.* (1997) also study the mismatch in Portuguese labour market using the 1991 Quadros de Pessoal data, where years of schooling variable is used to calculate the mismatch using several different approaches (explained previously in the objective measure part), and find that oth overeducation and undereducation exist in the market. Moreover, the phenomenon of overeducation should be a transitory situation (temporary), where workers accumulate human capital which eventually allows them to improve their job situation.

Furthermore, Green *et al.* (1999) offer a sensible explanation for overeducation; it is related to the fact that the quality and type of education acquired differs by institution and depends on the curriculum studied. Not all school graduates are equal in terms of their skills and productivity, or that educational human capital cannot be characterised as a homogenous stock. As such, worker productivity and earning will vary according to the quality and type of education obtained (grade, place of study, curriculum studied, *etc.*) and equally based on the demand for these different types of skills.

The Job Competition Model

The job competition model offers a demand side explanation for the existence of overeducation. This model is based on Thurow's (1975) book entitled 'Generating Inequality'. The model suggests that job characteristics may be the only factor to determine wages. The central element of the Job Competition Model is based around the observation (Thurow cites US surveys) that the majority of workplace skills are acquired through on-the-job training as opposed to formal education. Therefore, the labour market is not a bidding market for selling existing skills but a training market where training slots must be allocated to different workers. How these training slots are distributed across individuals depends on factors determining where individuals are located within a particular job queue and the distribution of jobs (training opportunities) in the economy. Once individuals reach the top of the queue and are allocated a job, their wages will be pre-determined by the characteristics of the job in question. Thus, the marginal product resides in the job rather than the individual's characteristics.

The Job Competition Model emphasises the importance of a person's relative position. Thurow postulates that were an individual to observe his/her neighbour participating in education, then under the human capital framework that individual would be less likely to participate in education, as the supply would be higher and the return less (Thurow, 1975). Slightly differently, under the Job Competition Model, the same individual would now be more likely to participate as education is a defensive need, necessary to protect their place in the queue. The larger the number of educated persons in the economy, the more imperative it becomes for individuals to invest in education.

The Job Competition Model provides a clear explanation for educational overinvestment as well as overeducation. The model is similar to the signalling framework in that preserving one's position motivates individual investments. In Spence's (1973) model, there is a limit to the amount of education in which an individual will invest (it is only based on the balance between wages and the cost of education). However, it is difficult to determine how a ceiling is reached within the Job Competition framework with respect to educational participation. This model provides a theoretical framework with which overeducation is entirely consistent. The implications of the Thurow's model are that wages will be wholly dependent upon required education and that the returns to education that are over and above that is required by the job (education surplus) will be zero (McGuinness, 2006).

In a similar respect, Muysken and Weel (1999) apply the job competition theory among skilled workers in European countries. They formalise the observation of increasing supply of skilled workers without rising wages. The study categorise the jobs based on skilled and unskilled jobs; and years of schooling is used as the proxy of education variable. They find that the results fit the facts for the Netherlands and other European countries quite well, since educational attainment had increased over the past decades whereas wages did not rise dramatically. The adjustment process went through on the one hand bumping down of skilled workers into unskilled jobs and on the other hand crowding out of unskilled workers into unemployment.

Moreover, Linsley (2005) studies overeducation in Australia labour market using the NLC survey. The study employs ISA and tests four of the key theories, *i.e.* the human capital, job competition, assignment and the career mobility theories. Various levels of education represent the education variable. The study concludes that job competition model is the best model to explain overeducation in Australia labour market. The model offers several implications: (1) overeducation persists and leads to less skilled workers being bumped down into low-skilled, low wage positions or crowded out of the labour market entirely; (2) a proportion of an individual's investment in education has limited productive benefit. Thus, reallocation of investment in education towards vocational education and training institutions which provide intermediate skills for individuals is necessary.

Assignment Models

The assignment models originate from Tinbergen's (1951) analysis of the determinants of income distribution. In this early model, there is a distribution of jobs (varying by some characteristics) and a different distribution of workers. Workers are affected by the mismatch between jobs and their own characteristics. Differences in wage rates arise which reconcile the distributions of jobs and workers by compensating workers for taking a job that does not match their characteristics (Sattinger, 2012).

Sattinger (1993) further asserts that there is an allocation problem in assigning heterogeneous workers to jobs, which differ in their complexity. The issue is that the frequency distributions on the demand and supply side are unlikely to match and education mismatches may be a persistent problem if the job structure is relatively unresponsive to changes in relative supplies of educated labour. Moreover, assignment models differ significantly from the Job Competition interpretation in that the models stress that choice of job or sector creates an intermediate step between an individual's characteristics and their wages, *i.e.* the job allocation process is not merely a lottery. Income maximisation guides workers to choose particular jobs over others. Subsequently, higher wages for workers with some characteristics play an important role in the economy rather than simply being rewards for the possession of particular characteristics. Workers find jobs in particular sector are not randomly distributed but are there based on the choices made to maximise their income or utility.

The central and crucial prediction arising from the literature is that in order to adequately explain changes in the distribution of wages, it must consider individuals and job characteristics. Therefore, overeducation is entirely consistent with the Assignment Interpretation suggesting that marginal products and wages will depend to some extent on both the individuals and the jobs; these models also imply that there is no reason to expect that wage rates will be wholly related to acquired schooling or other individual attributes (Human Capital theory), neither should it be expected that wage rates will be wholly related to the nature of the job (Job Competition Model). In short, match, overeducation and undereducation are determined by workers' and jobs' characteristics.

In terms of empirical evidence, Rigg *et al.* (1990) suggest that in the 1980s approximately one-quarter of UK employers had substituted university graduates for non-graduates, but that approximately one-third of the jobs where this substitution had occurred had not been upgraded in line with the higher educational requirements. Similarly, Pietro and Urwin (2006) study education mismatch in the Italian graduate labour market (university level) in 1997 and 2003, and find little evidence to support the assignment theory; individuals' pay is determined by both their human capital and the characteristics of their job. Instead, they identify a relatively weak wage effect arising from education mismatch associated with employers', as opposed to employees', perceptions of the job requirements. This is probably because employers have re-categorised jobs as requiring a degree, when they were previously filled by non-graduates, and many have not altered pay scales accordingly.

Career Mobility Theory

The theory of career mobility suggests several specific predictions especially related to the effect of schooling on wages and firm mobility. One of the predictions is the effect of schooling on the probability of being promoted from an occupation (within or across firms) will be higher if the returns to schooling are lower while one works in a specific occupation. Similarly, it will be rational for some individuals to spend a portion of their working careers in occupations that require a lower level of schooling than they have acquired. Thus, an individual may choose a job with a lower wage return to education but with a higher probability of promotion in the future than other available jobs with higher wage effects of schooling (Sicherman and Galor, 1990).

Furthermore, Sicherman (1991) applies this theory to explain overeducation, using the 1976 and 1978 waves of the Panel Study of Income Dynamics (PSID) with male heads of household, aged 18-60 years old as the sample. The education variable is presented by years of schooling, and occupation is in two digits occupational category. The study uses RM to analyse and ultimately concludes that overeducated people are willing to choose a job for which their educational levels are higher than needed in order to acquire skills or better opportunities to make a career upgrade in the future. However, Nielsen (2007) argues that this theory only takes effect if overeducated individuals indeed move to a higher level of job to fully utilise their educational qualifications. According to the literature, this theory could be unrealistic for two reasons. Firstly, there is a lack of empirical evidence. For instance, Sicherman (1991) finds that both overeducated people and undereducated people have a positive probability of promotion. Moreover, Büchel and Mertens (2004) point out that overeducated people remain in mismatched status within five years. Secondly, this theory fails to provide a rational explanation in terms of undereducation.

Signalling/Screening

Spence (1973) assumes that the labour market has an informational asymmetry. On the employers' side, they are not completely certain about true productivity of their employees when they make the employment decisions and even after hiring. There is a signal transfer mechanism passing the information about employees to the employers in order to identify the most productive and motivated workers for the firms. Education acts as an important signal in this mechanism, acting as a screening device representing some unobserved personal characteristics, such as problem-solving skills, communication skills and motivation. High educational levels signal high productivities. Individuals will keep on investing in education in order to distinguish themselves from others. Thus, investments in schooling are efficient from an individual's point of view but do not necessarily affect a worker's productivity.

This theory is not only able to help firms reduce the cost of hiring but is also beneficial for the final occupational distribution and placement in the labour market; by education signal, the most productive workers will be placed in the jobs that make them more productive. If high returns of investment in education remain, incentives for investment in education will last long (Tsang and Levin, 1985). Based on this theory, overeducation may be a persistent phenomenon, or in contrast to Freeman's argument (Freeman, 1976). In addition, signalling and screening are similar; the difference is that workers move first to choose their education level in order to signal their productivity to the employers. While in the screening theory, employers move first to decide the education level required for a job.

Green *et al.* (1999) sum up previous studies on overeducation and the signalling model in the UK. They highlight the findings from Rigg *et al.* (1990) that about 25 per cent employers in late 1980s had substituted graduates for non-graduates, and only about a third of these jobs had been upgraded in terms of content. Meanwhile, the Institute of Personnel and Development (IPD, 1997) conducted a survey of employers, recruitment firms, outplacement agencies and graduate careers offices. One in ten employers felt that they had a problem attracting too many over-qualified people to their job advertisements. This problem was even greater among medium and large firms, with one in four reporting a problem with over-qualified applicants. However, employers may use the term 'over-qualified' to politely indicate that a candidate is, in their view, too old or experienced for the job.

In terms of empirical evidence, Ordine and Rose (2009) study overeducation in Italian graduate labour market, using the data from the Italian Ministry of Education and the Italian commission for academic research evaluation. For the method, the study uses the probit model, and overeducation is defined by DSA. The study focuses on educational choices in a signalling setting in the presence of heterogeneous working ability imperfectly correlated with schooling costs. In light of the theory, the study shows that equilibria may occur with forward induction reasoning, in which individuals with different abilities acquire the same educational level. Meanwhile, the assumed strategic interaction between firms and individuals' choices considers explicitly the externality generated by low-ability individuals with low indirect costs who use education in order to signal the abilities that they do not have. As such, education mismatch derives from the emergence of pooling equilibria related to the dimension of the "ability effect" with respect to the "indirect costs effect" in human capital investment. When education quality is low, the latter effect may prevail, and it increases the probability of overeducation. The study further highlights that the larger the share of innovative firms is, the higher the quality of education must be in order to avoid overeducation. This is because low-ability individuals may see better job opportunities by acquiring higher education and they do not internalise the impact of their choices on the firms' behaviour. Moreover, empirically, the study concludes that overeducation in Italian graduates' labour market appears to be strongly determined by university quality and by other variables which characterise the individual's socioeconomic background. In addition, many other studies also use this model to estimate the wage premium rather than the determinant of overeducation.

Technological Change Theory

According to Oliveira *et al.* (2000), there are two possibilities for the existence of education mismatch, and one of them is technological change[36]. Rapid technological change may require school-provided skills or education higher than those already possessed by currently employed workers. Even in the presence of positive adjustment costs, better-educated workers cannot be made instantaneously. The employers and the employees could be locked into a situation of disequilibrium (at least in the short run) and hence pockets of undereducation would arise. In less flexible labour market settings (as is the case, in general, in European countries), these pockets might be persistent. At the same time, firms upgrade their hiring standards and the recently hired employees (those with higher educational qualifications than their older co-workers) are perceived to be overeducated.

Empirically testable hypotheses derived from the technology-based hypothesis contradict those of the human capital approach. Overeducated workers are those that the employer

[36] The other one is human capital.

wishes to retain and for whom the bulk of training and firm-specific investments is slotted; hence, their tenure is rewarded. Undereducated workers, on the other hand, are confined in dead-end jobs and should expect no such benefits from continuing with the same employer.

Oliveira *et al.* (2000) use an extensive dataset based on staff records (Quadros de Pessoal) collected in 1991 by the Portuguese Ministry of Employment from all business firms with more than one paid employee; and apply a decomposed Mincer wage equation. Oliveira *et al.* test two competing hypotheses; the first hypothesis, undereducation is the outcome of a process in which market-acquired capital substitutes for insufficient school-supplied qualifications, whereas overeducation is associated with excess schooling but short tenure and job experience. And the second one considers upon changes in the technology of production and marketing to explain why some workers end up as inadequately educated for the tasks that they perform, while at the same time, others (holding identical jobs but more schooling) are perceived to be overeducated. They find employers tend to value and prize overeducation and at the same time they penalize undereducation. With prolonged tenure, overeducated workers are granted an ascending path of their relative earnings, while undereducated workers will see their relative position eroded. Furthermore, there is an indication that overeducation or undereducation in the Portuguese economy occurs due to the role of technological change rather than human capital. This is characterised by intensive efforts to promote economic growth, modernisation of the industrial structure and the upgrade of educational qualifications in the last decade.

Carnevale and Rose (2013) also study undereducation in the US using the statistical framework developed by Autor *et al.* (2008) while the supply and demand are estimated based on Goldin and Katz's (2008) data. They argue that undereducation increases in the US because of under-producing and high demand of workers with higher education qualifications. If qualified workers are in short supply relative to the employers' demand for them, the rational response on the part of employer's is to bid up wages for the workers they want. As a result, the large and growing gap between the wages of workers with different education attainment grows even wider.

4.2.3 Determinants of Mismatch

Referring to the compilation of selected related studies (Appendix XI), the present study classifies the determinants of mismatch into some categories: (1) personal characteristics such as gender, ethnicity and age; (2) household characteristics such as number of young children in the household; (3) work related and firm size, including job status, tenure, sector, dummies of industry and dummies of firm size; and (4) residence, including urban/rural and capital/non-capital province. In addition, other macroeconomics variables can be considered in the model, such as GDP per capita as well as the share of manufacturing and regional unemployment. The analysis then focuses on the first four, since some macroeconomic variables are similar to those categories, for instance, the share of manufacturing which is represented by dummies of industry (including manufacturing) in the model.

Personal Characteristics

In terms of personal traits, variables which have received a large amount of attention in recent literature are sex, marital status, ethnicity, age, age squared and subject of study.

Firstly, a clear picture on the direction and significance of the effect of sex has yet to emerge. So far, many studies find that male employees face a slightly higher overeducation risk (European Commission, 2012) than female employees. This is also confirmed by Yin (2016); males were more prone to be overeducated than females in China during 1989-2009 period, as China has the patrilineal system is key to traditional Chinese family and gender values (Hu and Scott, 2014), males is perceived to have more responsibility to support the family life than females, and thus more likely to accept jobs that require lower educational level than their own. In contrast, *Clark et al.* (2017) analyses overeducation in the US, using the NLSY79 data combined with the CPS from 1982 to 1994 with the probit model and finds that females were about 5 to 13 per cent more likely to be overeducated than males. However, Chua and Chun (2016) find that females are more likely to be overeducated, while males are slightly more likely to be matched or undereducated. Other studies have also found that gender has insignificant effect on matches (Büchel and Pollmann-Schult, 2001; Battu and Sloane, 2002; Groot and Brink, 2003; Frenette, 2004; Green and McIntosh, 2007; Capsada-Munsech, 2015).

In terms of marital status, the theory of Differential Overqualification developed by Frank (1978) states that overeducation frequently occurs among females, particularly the married ones. The theory hypothesizes that married individuals, especially women, face a significantly higher probability of being overeducated, which is a consequence of matching problems. In traditional gender role model settings where the couple's priority is the job match for the husband, the husband acts as a first-mover, *i.e.* he performs his job search first. After he has found a match, the wife will conduct her job search. However, due to the co-location restriction, she can do that merely within a much smaller market area. The likelihood of finding a job adequate to her qualification level is therefore much lower for her than for her husband, thus explaining a striking incidence of overeducation among married females (Boll *et al.*, 2016). Having said that, McGoldrick and Robst (1996) reject the hypothesis of differential overqualification; instead, it appears for them that the larger number of vacancies in large labour markets are offset by a larger number of job searchers. Likewise, Battu and Sloane (2002) find that marital status have no significant effect on education mismatch.

With respect to ethnicity, Battu and Sloane (2002) assert that there are no studies that explicitly focus on mismatch amongst ethnic minorities. However, the limited research on ethnic group and overeducation finds that overeducation incidence is greater for non-whites (minorities). An argument specific to non-whites is simply that of discrimination. Theoretically, the same argument can be applied for females and different religion. Yet, empirically gender have no significant effect on the probability of being matched or otherwise (Battu and Sloane, 2002). If non-whites find it more difficult to acquire any jobs they may well be more likely to take a job that is not proportional to their qualifications, so that a higher number of non-whites will end up being over-educated. Battu and Sloane further focus on overeducation among ethnic minorities in Britain, using the Fourth National Survey of Ethnic Minorities (FNSEM) in 1993/1994. The ethnic population is composed of six groupings (Caribbean, Indian, Pakistani, African-Asian, Bangladeshi, and Chinese), while education attainment covers non-educational, O-level, A-level and university degrees. The study uses a multinomial logit model to investigate the determinants of over and undereducation. The dependent variable is earnings, which refer to usual gross pay from the sample's main job including overtime and bonuses before any deductions[37]. The study finds that only African-Asians have a significantly greater likelihood of being mismatched, relative to the omitted category of Indians. Using interaction terms, the study concludes that African-Asians who are born in the UK and who have foreign qualifications are less likely to be overeducated. By way of contrast, Pakistani and Bangladeshi workers with foreign qualifications have a higher probability of being overeducated.

Turning to age, the literature (such as Boll *et al.*, 2016) asserts that age and overeducation have a negative relationship, suggesting that there is a risk-reducing effect of age. Similarly, Flisi *et al.* (2014) affirm that older age category implies lower probability of being severely mismatched rather than matched in all countries. Regarding overeducation, common results are found for older age groups, since in almost all countries age groups 55+ and 45-54 have lower probabilities of being overeducated than age group 35-44. But for younger age groups, there are some relevant differences by country: age group 25-34 is less likely to be overeducated than age group 35-44. The effect of the square of age is negative, implying that there are diminishing returns in age. The quadratic function of age variable could capture the fact that the marginal effect of education mismatch declines over time.

Other variables commonly included in previous empirical studies are subject of study (Silles and Dolton (2002), McGuinness (2006), and Boll *et al.* (2016)). Boll *et al.* study the overeducation in 25 EU countries using the European Labour Force Survey (EU-LFS) data; and emphasise that high-skilled workers from the fields of agriculture, veterinary and services are much more frequently overeducated than high-skilled workers from the fields of teaching, education and health, and welfare (Ordine and Rose, 2009).

In addition to determinant variables, education mismatch may be influenced by the other skills dimension, such as lifelong learning, on-the-job training. Vocational lifelong learning is the responsibility of the Ministry of Manpower that conducted in both public and private employment training centres. In 2014, there were 1,555 private employment training centres in Indonesia–almost five times the number of public employment training centres (Lee Kuan Yew School of Public Policy and Microsoft, 2016). In addition, for

[37] Earning is in bands. The study finds that earnings for non-whites overall are significantly lower, with Bangladeshis having the lowest earnings and African-Asians and the Chinese displaying parity with Whites. It is worth noting that at one point in history, an Indian master degree was considered in the UK equal to a UK undergraduate degree.

lifelong learning, the present study identifies at least two programmes that the government have launched: improving adult literacy and digital business training; according to the Ministry of Education and Culture (2013) around 3 million people participate in literacy programmes in Indonesia. A particular emphasis is placed on increasing women's literacy levels, combining more generic life skills with literacy courses. For on-the-job training, the government encourages industry or companies to conduct competency-based job training for workers and prospective workers (TNP2K, 2015). Those variables are available in the IFLS data. There are some questions related to on-the-job training and life-long learning: have you ever received any training from your employer? What kind of training did you receive in the last 12 months? The answers are computer, language, technical training, teamwork, leadership, and others.

Household Characteristics

In terms of household characteristics, most of the relevant literature has focused on the presence of children as a determinant of overeducation. Boll *et al.*, (2016) find that the coefficients of children are insignificant in general; children of any number and age composition do not affect overeducation risk for male workers. Yet, the interaction terms with gender are relevant for female workers. More specifically, having an additional child below the age of six is predicted to reduce the overeducation probability significantly for high-skilled female workers. For medium-skilled workers, the risk-reducing effect of small children is of lower magnitude and is only weakly significant. A reason could be because medium-skilled workers are, on average, expected to be less wealthy than the high-skilled ones, which could force them to accept barely adequate jobs when living with children.

Meanwhile, Sloane *et al.* (1999) assert that the presence of young children in the household poses different effects across gender, as younger children reduce overeducation for males and raise overeducation for females. Hence, females with children are forced to make more compromises in the labour market. Empirically, raising the number of children between the ages of 0 and 2 from zero to one reduces the probability of being overeducated for males by 7.38 per cent and raises the probability of being overeducated for females by 17.32 per cent. Sloane *et al* use subjective measure based on a question in the Social Change and Economic Life Initiative (SCELI), which is funded by the Economic and Social Research Council (ESRC). The survey covered six British local labour markets between 1986 and 1987.

Other household-related variables that could affect mismatch determinant are number of adults with unemployment status in the household, number of inactive people, and number of older people.

In contrast, Dolton and Silles (2002) conduct a research based on the Newcastle Alumni Survey, collected at the University of Newcastle-upon-Tyne in 1998, and find that there was no measurable effect between children and marital status on overeducation in the UK labour market (in particular university graduates). A possible explanation is that the sample is not large enough to allow a meaningful interaction between family commitments and gender.

Work Related and Firm Size

In work related, another potential variable is working experience. Both the Human Capital Theory and the Job Competition Theory confirm that there is a negative relationship between work experience and overeducation risk. Moreover, the Career Mobility Theory asserts that the longer a worker stays in a firm, the higher is the likelihood of advancement into better positions with higher skill requirements and thus a lower overeducation risk. Meanwhile, Groot (1996) finds a positive relationship between experience and overeducation probability, since low-productive workers receive fewer job offers and therefore tend to remain stuck in bad matches which under-utilize their skills. The effect of tenure is similar to experience, both of them significantly reduce the risk of overeducation (Büchel and Pollmann-Schult, 2004). This is because individuals accumulate human capital by working. Furthermore, human capital can be dichotomised into general and firm-specific human capitals. Moreover, the quadratic function of experience and tenure (squared) variable could capture the fact that on-the-job training investments decline over time in a standard lifecycle human capital model, as explained in Chapter 3.2.2.

In terms of job status, some studies use full-time and part-time (less than 30 hours a week) as job status variable (Lindley and Machin, 2016) which is commonly determined by the number of working hours a week. Other studies prefer to use the alternative proxy *i.e.* working hours per week or per year (Clark *et al.*, 2012, and Boll *et al.*, 2016). In particular, Boll *et al.* (2016) find that working hours have a negative and significant coefficient, which implies that workers with more working hours are less likely to become overeducated, especially in high-skilled jobs. This is because jobs with longer working time can create better opportunities for training participation and advancement, thereby improving the match quality over time. Similarly, Frank (1978) and Ofek and Merrill (1997) add that part-time work leads to a higher probability of being overeducated. Sloane *et al.* (1999) also assert that being a part-time worker reduces the probability of being undereducated and increases the probability of being overeducated. They argue that a part-time work for the overeducated may simply facilitate job search, thereby representing a short-lived mismatch as part of a longer-run career development path. Meanwhile, the European Commission (2012) does not detect any significant associations between education mismatch and working hours or job status. Morano (2014) also finds that part-time and temporary employments both lead to higher overeducation.

Turning to sectors, Dolton and Vignoles (2000) study overeducation in the UK using the 1980 National Survey of Graduates and Diplomats. They put forward that education mismatch, particularly overeducation, is found in broadly equal proportions in both the public and private sectors. However, a higher proportion of those working in government administration were specifically overeducated in 1986. Similarly, Ortiz (2010) studies overeducation in France, Italy, and Spain, using The European Community Household Panel from 1999 to 2001. The method used to calculate overeducation is RM. The Multinomial Logit (MNL) is also applied to estimate the determinants. The study finds that working in the public sector increases the likelihood of being overeducated, relative to working in the private sector. The main reason is working in the public sector is generally considered more secure. In contrast, Yin (2016) finds that workers in the private and collective sectors in China are more likely to be overeducated than individuals in the government sector. This is likely due to the government sector having higher wages, stable working environment and attractive welfare, which makes it easier to hire matched workers. Another possible reason is because individuals may find jobs in the private sector only temporarily to gain experience in order to find better matched and stable jobs in the government sector later.

Regarding occupation, Dolton and Silles (2002) study the determinants of graduate overeducation in the UK using data from the Newcastle Alumni Survey. They employ cohort effect to analyse the effect of initial overeducation on the probability of being overeducated in the future. Most importantly, the study includes occupation (manager, professional and associate professor, with the base group being other occupations). The study finds that graduates in professional, associate professional and managerial occupations have a greater propensity to be in graduate level jobs than those in the base group. In similar vein, Morano (2014) analyses the determinants of overeducation in Italy using the Continuous Labour Force Survey (Rilevazione Continua delle Forze di Lavoro). The study uses some occupation categories in the model: director, manager, blue-collar worker and trainer, with the base group being clerical jobs. The study finds that these occupations have negative and significant coefficients; with the exception of blue-collar workers who have a positive and significant coefficient. Negative and significant coefficients indicate that the probability of being overeducated is lower for these categories than for the base group (clerical jobs).

Another possible variable to influence education mismatch is industry, as Allen (2016) asserts that occupation mismatch in Indonesia tends to be associated with the low education levels of production workers and agricultural laborers. On the other hand, a large number of clerks are over-qualified for their jobs. Undereducation is also a challenge in higher-level occupations. The high levels of under-qualification and lower levels of over-qualification point towards an issue of skill shortages. Meanwhile, Morano (2014) finds that workers in the service sector are less likely to be overeducated than those in the agricultural or industry sectors. These results can be interpreted by the different nature of employment in the three economic sectors and the relatively less skilled nature of the jobs in the agriculture or industry sectors with respect to service work.

Furthermore, firm sizes could also determine education mismatch, as indicated in previous empirical studies. For instance, Dolton and Vignoles (2000) find that overeducation was the highest amongst those who work for small firms (of less than 20 people), although generally the incidence of overeducation does not decrease linearly with firm size. Interestingly, more than 70 per cent of the graduates who are overeducated and work in small firms claim to require no qualifications for their job. It may be that the lack of benchmark jobs and formal qualifications causes a higher incidence of overeducation to be recorded amongst graduates who work in small firms. Morano (2014) also finds that overeducation decreases as firm size increases, which is consistent with the idea that bigger firms have more accurate recruitment techniques which reduce the risk of hiring a worker who does not match the educational requirements associated to the vacancy (Dolton and Silles, 2001). Moreover, there is a wider range of positions that enables the management to internally relocate workers in case of mismatch in big firms. Yin (2016) also adds that workers in large firms with between 20 and 100 employees and in firms with fewer than 20 employees are less likely to be overeducated than workers in firms with more than 100 employees. This could be due to employees in large firms having more opportunities to be promoted, having better career prospects and receiving more fringe benefits compared to those who work in a small company. Thus, overqualified workers may voluntarily choose to stay in large firms due to the consideration of the above benefits. This finding is in contrast to Morano (2014), implying that firm size definition, country and time period may lead to different conclusions.

Another alternative variable is job contract length. Boll *et al.* (2016) consider the incidence of overeducation is strongly related both to job type (includes contract length) and firm characteristics in 25 European Countries. The study argues that people with fixed-term contracts are more likely to work in positions for which they are overeducated than people with permanent contracts. This is due to the transitory nature of fixed-term jobs; workers are less concerned about qualification levels, as they tend to view these matches as mere temporary solutions on their way to more favourable permanent positions. Similarly, Green and McIntosh (2007) and Ortiz (2010) also identify evidence for a significantly lower overeducation risk among workers in permanent positions.

In addition to determinant variables, Chua and Chun (2016) emphasise that labour markets in developing countries are unique for their large shares of informal sector employment. This sector is comprised largely by either microenterprises or menial wage work, where high level skill or training is not required. The study also adds that the formal sector has better matches than the informal sector, even the informal salaried and self-employed sectors. In fact, overeducation is particularly severe among self-employed workers, which plausibly explained by the preponderance of small businesses with low skill requirements.

Area

In terms of residence or area category, urban/rural is used in some studies, such as Clark *et al.* (2013) study the relationship between education mismatch and urban area based on the NLSY 1979 combined with the pooled 1989-1991 waves of the CPS. They found that workers in urban areas in general have a higher risk of being overeducated than workers in rural areas, which is surprising. One would expect cities and areas with low unemployment rates to have a labour market with relatively lower levels of search frictions, making it easier for individuals to sort into an occupation with a required level of education that matches theirs. However, they empirically find insignificant coefficient of urban areas in the US.

Likewise, Clark *et al.* also use regional or unemployment rate variable as a mismatch determinant. The study finds that individuals who live in urban areas are more likely to be overeducated, while individuals living in regions with low unemployment rates are less likely to be overeducated. A possible argument is that cities and areas with low unemployment rates have a labour market with relatively lower levels of search frictions. As such, it becomes easier for individuals to sort into an occupation with a required level of education that matches their educational qualifications.

In the developing countries, the trend could be different. Yin (2016) argue that urban workers in China have a higher average educational level and better quality of education than total workers. Yet, job vacancies in urban labour market are more abundant than in rural areas, competition is sharper in urban areas than in rural areas. Under the fierce competition in the urban labour market, urban workers may work in a job that is lower than education level in order to make a living and thus they have higher probability of being overeducated than rural workers. This is consistent with Allen (2016) that most new jobs are generated in urban areas in Indonesia.

Quang and Tran-Nam (2019) study mismatch and earning in Vietnam using Duncan and Hoffman (1981) model and the Labour Force Survey as the data. They argue disaggregation by regions of settlement does not show significant difference in the distribution of mismatch. However, the situation in rural areas seems to be better, evidenced by higher well-matched rate and lower under-education rate. One of possible reasons is people move from rural to urban areas to find employment, which, thus, raises the incidence of under-education in the city.

In addition, the Harris Todaro model (1970) can explain rural-urban effect on internal migration. Although the model does not directly explain the effect on education mismatch, it gives an explanation that rural-to-urban migration still exists despite high levels of urban unemployment. The model argues that the labour market in developing countries is not clear. In short, the model is as follows: the urban minimum wage is set to be higher than the wage rate paid to rural labour which is valued at its marginal product. This results in a wage differential between the two sectors. Rural workers have an incentive to migrate to the urban areas despite of urban unemployment, because of the potential of higher earnings in the urban sector. Such migration will continue as long as there is a possibility for migrants to increase their income by moving to a city. Some migrants will have arranged employment before leaving the countryside. Others will begin searching for employment only once they have arrived at their destination. Some will necessarily join the pool of urban unemployed. However, even in that case, the presence of the migrant in the city may increase his chance of finding urban employment there at a later date. This explains why there is a continuous flow of migrants observed in developing countries despite of the high urban unemployment rates (Bahns, 2005).

4.2.4 Aggregate Trends and Comparison of the Prevalence of Education Mismatch

Empirical studies find various results on the change of education mismatch depending on the methods and data used. Whereas some studies find that the trend increases (Nazara and Safuan (2005), Mehta *et al.* (2011), McGuinness *et al.* (2017)); some others find it decreasing (Yin, 2016); and even remaining stable (McGuinness, 2006). In particular, Nazara and Safuan (2005) find that the trend of undereducation in Indonesia decreases.

McGuinness (2006) tries to plot in the graphs based on subjective and objective measures to determine whether the overeducation phenomenon may be becoming more important over time. Based on subjective measures, there are no indications that the incidence of overeducation has been rising over time; in fact, fitting a linear time trend to the observations is suggestive of a slight decrease instead. Nevertheless, given the problematic nature of the data, it would not be wise to attach too heavy a weight to the very slight negative slope of the best-fit line. Slightly differently, objective measures have a positive, sloped trend line. Thus, the study argues that it is probably reasonable to conclude, on the basis of the graphical and tabular evidence, that the incidence of overeducation has remained relatively stable during 1971 – 2000 periods.

Furthermore, McGuinness *et al.* (2017) analyse the patterns in overeducation between countries using a specifically designed panel dataset constructed from the quarterly

Labour Force Surveys of 28 EU countries over a twelve to fifteen-year period. Mcguinness *et al.* use the Barro regressions[38] (Barro, 1997) to analyse the relationship between the initial level and the growth rate of overeducation. The study concludes that overeducation is consistently rising across all European countries; in fact, there is a positive trend in fewer than half the countries in the sample which shows overeducation remaining either constant or falling in most cases. Moreover, overeducation is higher among females in the vast majority of these countries. In the more aggregated level, the average trend in overeducation across all 28 countries appears to be relatively stable over the period of 2003 - 2013; though substantial differences do exist depending on the geographical country block. Overeducation rates tend to be the highest and most volatile over time in peripheral European countries, while overeducation in central European countries tends to be lower and appears to follow a somewhat cyclical pattern. Overeducation is consistently the lowest and stable over time in eastern European countries. There is also an on-going convergence in overeducation, where countries with the lowest initial values of overeducation tend to experience the highest growth rates over time. Finally, in terms of the factors driving cross-country differentials, factors relating to both the composition and level of labour demand, labour supply and the structure of educational provision all appear important.

In Asian and other developing countries, Mehta *et al.* (2011) study overeducation in unskilled jobs in several developing countries (India, Mexico, The Philippines and Thailand) between 1990s and 2000s. Overeducation is determined by RM method and by comparing the mean and the mode. The study find that overeducation increased slightly in the 1990s while undereducation decreased slightly under the mean criterion. The opposite result is found under the mode criterion, where overeducation declined sharply and undereducation rose substantially. Nonetheless, the result shows that a job's mean and modal years of schooling are poor proxies for required education. Moreover, it also indicates that only 25 per cent of the overeducated workers under the unskilled jobs test are classified as overeducated by the mean method. In other words, 75 per cent of workers who they identify as earning low returns relative to their education are not classified as overeducated by the mean method, because their education backgrounds are not atypical.

---

[38] The Barro model is used to examine the relationship between the growth rate of overeducation and the initial level of overeducation using as regression model, as formulated by: $\frac{\ln Ov(t) - \ln Ov(0)}{t} = \beta_0 + \beta_1 \ln Ov(0) + \varepsilon$

Meanwhile, Yin (2016) adds that the marginal effects of overeducation in China for seven waves are negative, possibly indicating that the years after 1989 are less likely to experience overeducation compared to the year 1989. This decrease could be because the average educational level of the labour force is increasing, and the increasing degree of economic openness and the accelerated growth of the private sector may provide more opportunities from which the workers can choose.

Handle (2017) analyses education and skills mismatch in developing countries, using The World Bank's Skills Measurement Program. 12 countries which were observed are: Ghana, Kenya, China – Yunnan, Lao, Sri Lanka, Vietnam, Bolivia, Colombia, Armenia, Georgia, Macedonia, and Ukraine. The data were collected between March 2012 and August 2017. The sample was random but was in working age of between 15 – 64 years old. Handle then highlights several issues in developing countries which affected the mismatch: a very high rate of informality, self-employment, and micro-firms (around 55-80 per cent of the total firms). Very low employment rates among the working age population (around 33 – 55 per cent of the total population) creates other issues such as selection issues and unemployment or inactivity. In turn, these issues reflect a very weak job market and low job generation as well as gender dynamic. Handle also argues that the mismatch likely occurs due to some drivers, such as: (1) labour market friction (imperfect information); (2) transitory business cycle, for example increasing unemployment and job seekers choosing any jobs available without considering their education level and background; (3) life cycle stage, particularly the youth; (4) work/family preferences, for instance women with young children; and (5) social exclusion, such as minority ethnicities or immigrants. Besides, education mismatch prevalence could also reflect problems with education (such as quality) and with the job market (such as low employment rate and low investment) in the country. This finding is consistent with the Harris-Todaro model, as explained in Section 4.2.3.

In Indonesia, Nazara and Safuan (2005) study the overeducation in the formal sector in the Indonesian labour market, using Sakernas data from 1996, 1999 and 2002. Overeducation and undereducation are defined as the deviation from the mean of years of schooling. Furthermore, occupation is divided into 9 groups: professional, management, administrative, sales, labour service establishment, agriculture, production workers and labour, transportation operator, and moving equipment and unskilled workers. The study reveals that there is an indication that overeducation exists in

Indonesia. Moreover, the study estimates that undereducated, matched and overeducated workers in 1999 were 16.8 per cent, 56.5 per cent, and 26.7 per cent from the sample, respectively. The share of undereducation then decreased to 9.13 per cent in 2002; in contrast, the share of overeducation increased to 34.7 per cent. The study argues that overeducated workers exist probably due to limited choices resulting in a very competitive labour market for highly educated people whilst the job-search cost is relatively high. Alternatively, another possible explanation is that this result only reflects the distribution of ability. Wajdi *et al.* (2017) also study the urbanisation in Indonesia. Even though it is not directly related to overeducation, their findings may support the argument that migration to DKI Jakarta (the capital city) and other metropolitan areas is most likely undertaken for better education or job prospects, since most jobs for highly educated people are more available in the metropolitan areas.

Similarly, Allen (2016) measures the mismatch using the International Standard Classification of Occupations (ISCO) and the International Standard Classification of Education (ISCED). This measure of mismatch divides major occupational groups (1-digit ISCO levels) into four sub-groups and assigns a level of education to each occupational group in accordance with the ISCED. Workers in a particular group who have the assigned level of education are considered well-matched. Those who have a higher (lower) level of education are considered overeducated (undereducated). The study finds that 51.5 per cent of workers are undereducated, 40.0 per cent are matched and 8.5 per cent are over-educated for their occupations, as shown in Figure 4.1.

In particular, occupation mismatch (overeducation) in Indonesia tends to be associated with the low education levels of production workers and agriculture labourers, as well as a large number of clerks who are overeducated for their jobs. Meanwhile, undereducated is also an issue in higher level occupations, such as legislators, senior officials, and managers (Figure 4.1). Furthermore, the large proportion of undereducated workers could be a reason for the slow labour productivity growth and slow transition to high value activities throughout the economy, as well as for the prevalent skill and/or education shortages to occur in Indonesia.

Moreover, education becomes the most important factor for career progression; production workers with post-secondary education are likely to have upward career mobility into more technical or managerial occupations over a 12-month period. Meanwhile, once workers with tertiary education move into professional and technical occupations, they are likely to still be working in such occupations 12 months later. Thus, workers with university qualifications tend to have a higher incidence of long-term overeducation, which could be explained by jobs rationed by queuing[39]. Allen (2016) also further asserts that limited education attainment may tend to act as a barrier to career progression, for instance, workers with junior high school or lower qualifications are likely to shift between working as production workers and as agriculture labours throughout the year. However, these workers find it more difficult to climb the career ladder.

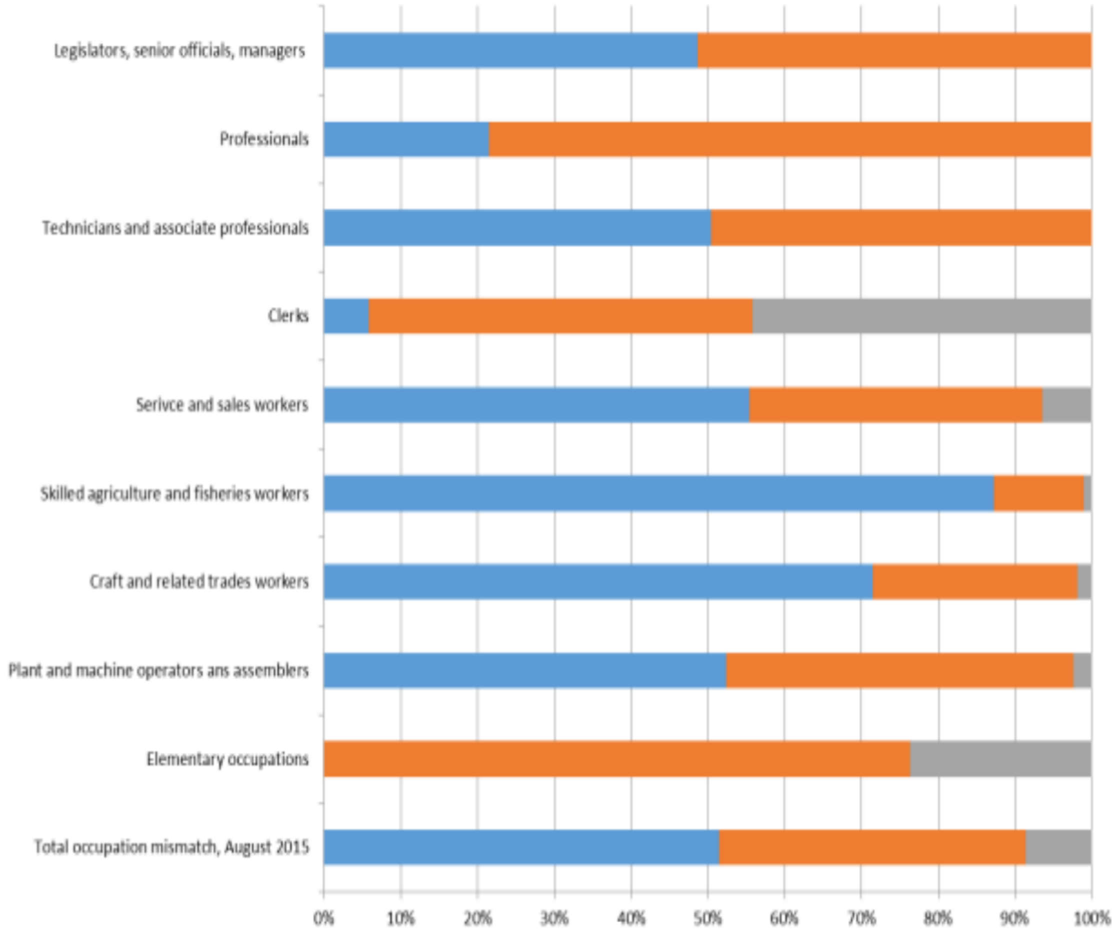


Figure 4.1: Education Mismatch in Indonesia, August 2015
Source: Allen (2016).
Note: Blue is undereducated, orange is matched, and grey is overeducated workers.

[39] Jobs rationed refer to a system of shared beliefs about who should have access to the job market. Given the limited supply of jobs, a system of norms has developed about how work should be distributed. Thus, certain individuals are encouraged to consider themselves unsuitable candidates for employment under existing conditions of job scarcity (Furstenberg, 1975).

ILO (2017) also calculates the undereducation and overeducation levels in Indonesia by using RM method (comparing the mean of occupation group and its standard deviation). The data used are SAKERNAS from 2006, 2009 and 2016. ILO finds that there was an increasing trend of undereducation from around 10 per cent in 2000 to around 17 per cent in 2016. Meanwhile, overeducation trend decreased from around 27 per cent in 2006 to 19 per cent in 2009; though it then became relatively stagnant at around 19.2 per cent in 2016. ILO also reveals some interesting trends during these periods: male workers tend to have lower education level than female workers; higher undereducation rate occurs in urban areas; older generation experiences being undereducated which could also indicate improvement of education achievement; and around 25 per cent of people aged 15-34 years old tend to experience overeducation.

4.3 Method and Data

4.3.1 Method

Marginal effects (ME)

Besides the MNL estimation, the present study also provides a marginal effect calculation for each dependent variable. It is worth noting that those coefficients of the MNL are different from the marginal effect. Multinomial logit coefficients can only be interpreted in terms of relative probabilities (this will be discussed in the next part) whereas the marginal effect calculation is needed to reach conclusions about the actual probabilities. Furthermore, marginal effects can be an informative means for summarising how changes in a response are related to changes in a covariate. As Cameron and Trivedi (2010) point out, a marginal effect, or partial effect, most often measures the effect on the conditional mean of y of a change in one of the regressors, for example; x_j. In addition, the marginal effect calculation is based on the first order derivatives; for interaction (age squared and tenure squared) between two variables, the second order derivative is then required.

The multinomial logit (MNL) model

Most studies on overeducation use binomial logistic regression (logit) model to analyse the determinants of overeducation. Some other studies use multinomial logit model to extend their analysis and to investigate the determinants of both overeducation and undereducation. MNL is a simple extension of the binomial logistic regression model, allowing for more than two categories of the dependent or outcome variables. It is used to predict categorical placement in or the probability of category membership on a dependent variable based on multiple independent variables. The independent variables can either be dichotomous (binary) or continuous (interval or ratio in scale). MNL uses maximum likelihood estimation to evaluate the probability of categorical membership, similar to binary logistic regression. MNL is also often considered an attractive analysis because it does not assume normality, linearity, or homoscedasticity (Starkweather and Moske, 2011). MNL specification is tractable and simple to estimate. And the most important thing is the independent errors of MNL force an assumption called the independence of irrelevant alternatives (IIA) assumption. As Dow and Endersby (2004) asserted the idea of IIA is that if a chooser is comparing two alternatives according to a preference relationship, the ordinal ranking of these alternatives should not be affected by the addition or subtraction of other alternatives from the choice set. Thus, the IIA property is a minimal condition for logical consistency. The probabilistic analogue imposed by MNL, strengthens this by requiring that the odds ratio of choosing any two alternatives be independent of the addition or subtraction of other alternatives from the choice set. Specifically, the ratio of choice probabilities for any two alternatives does not depend on the characteristics of any of the other alternatives. This is consistent with the context of education mismatch; a worker can be either matched or mismatched (undereducated/overeducated) without any influence from the other alternatives. It is worth noting that the IIA is a logical property of decision-making, not a statistical property such as consistency and unbiasedness.

Moreover, most of the studies which analyse education mismatch adopt MNL, for instance Chevalier (2007), Battu and Sloane (2002), Kiker *et al.* (1997), Flisi *et al.* (2014) and Diem (2015). Thus, the multinomial logit (MNL) model in the present analysis is as follows:

$$M_{i,t} = \beta_{0,t} + \Sigma\beta_{1,n,t}\mathrm{P}_{i,n,t} + \Sigma\beta_{2,n,t}\mathrm{HH}_{i,n,t} + \Sigma\beta_{3,n,t}\mathrm{F}_{i,n,t} + \Sigma\beta_{4,n,t}\mathrm{A}_{i,n,t} + \varepsilon_{i,t} \quad (4.1),$$

where:

$M_{i,t}$ is Match category (the dependent variable), consisting of 1 to 3, where: 1 is overeducated, 2 is matched and 3 is undereducated. Σ represents set or vector of explanatory variables; $P_{i,n,t}$ is personal characteristics ($1 \ldots n$ number of personal characteristic variables); $HH_{i,n,t}$ is household characteristics ($1\ldots n$ variables); $F_{i,n,t}$ is work-related and firm size variable ($1 \ldots n$ no of variables); $A_{i,n,t}$ is residence or area dummy variables (1…n number of residence dummies). i is individual (1…I); and t is at time t (2000 or 2014). It is also worth noting that the coefficients in equation 4.1 are not the same in terms of notation as in 4.2. The model is a probabilistic model, so the coefficients in 4.1 are interpreted as relative probabilities.

The baseline category of this analysis is the matched category; thus, the interpretation of the variables would be the likelihood of being in either one of the remaining two mismatch categories (overeducated and undereducated), compared to being matched. The variables are considered to be statistically significant at 1 per cent, 5 per cent or 10 per cent significance level.

The multinomial probit (MNP) model

Another alternative method is the multinomial probit (MNP) models, as used by Berlingieri and Zierahn (2014). The main different between MNL and MNP is assumptions about the probability functions. Technically, MNL and MNP are very similar; they differ only in the distribution of the error terms and each model has its own advantages and disadvantages. The advantages of MNL are that the specification is tractable, simple and faster in estimating the models. However, MNL imposes the restrictive assumption that choices are independent across alternatives.

These disadvantages of the MNL's can be solved by using MNP. The primary advantage of MNP relative to MNL centres on the IIA property. MNP has errors which are not necessarily independent, and these errors are distributed by a multivariate normal distribution. MNP does not assume IIA. On the other hand, MNP imparts a number of potentially serious problems. These are sufficiently difficult to detect that, in the absence of investing exceptional effort in model diagnostics, researchers are justified in using the MNL specification. The most important problem is that even formally identified MNP specifications are often weakly identified in application. This is serious because weak identification is difficult to diagnose and may lead to plausible, yet arbitrary or misleading inferences. The MNP presents a difficult maximum likelihood optimization problem that sometimes fails to converge at a global optimum or produces parameter estimates that are sufficiently imprecise as to make statistical inferences suspect. Except for cases of profound misspecification, the logit likelihood will optimize at its global maximum and is not prone to optimization errors (Dow and Endersby, 2004).

Considering those advantages and disadvantages, the present study prefers the MNL as the main model of the present study since MNL is appropriate in this context, IIA property can be fulfilled. Some studies even find that logit estimation (MNL) performs as well or even better than MNP, such as Dow and Endersby (2004). Yet, this study will also perform MNP as a sensitivity analysis and the results of both models will be compared.

4.3.2 Measures

Mismatch Determination

The present study analyses the education mismatch in Indonesia in 2000 and 2014 using the multinomial (MNL) model. The model is adopted from Battu and Sloane (2002) to estimate two sets of coefficients: β_1 which represents the overeducation coefficient and β_3 which represents the undereducation coefficient. The probability p_{nm} of individual n being overeducated (m=1) or undereducated (m=3) is conditional on a vector of characteristics K_i (vector of explanatory variable of mismatch determinants). And the probability of individual n being in the over or undereducated group m (relative to the probability of being in the default group 2 (matched) is given by:

$$\frac{p_{im}}{p_{i2}} = \exp\left[K_i'(\beta_m - \beta_2)\right] \text{ for m = 1,3} \tag{4.2}$$

with normalisation of β_2 to equal 0. To permit identification of the model, the probabilities are:

$$p_{i2} = \frac{1}{[1+\sum_{m=1,3}\exp[K_i'\beta_m]]} \text{ for m=2} \tag{4.3}$$

$$p_{im} = \frac{\exp[K_i'\beta_m]}{[1+\sum_{m=1,3}\exp[K_i'\beta_m]]} \text{ for m=1,3} \tag{4.4}$$

Educational match (M) is defined based on the difference between the workers' highest educational attainment and the mean of education attainment in the same occupation (as the proxy of job requirement). Thus, the mismatch formula is as follows:

$$M = Education\ attainment - (mode\ of\ education\ within\ occupation\ \pm\ its\ standard\ deviation) \tag{4.5}$$

The formula is adapted from Hartog (1997). M has three categories: 1 is overeducated, 2 is matched, and 3 is undereducated. Individuals are defined as being correctly matched (category 2) to their occupation if their own years of education or education levels are within plus or minus one standard deviation[40] of the mode for their occupation. Alternatively, the mode can be replaced by the mean to calculate match (this will be further elaborated in the sensitivity part[41]).

Furthermore, an alternative measure of overeducation is provided as defined by Chevalier (2000), who separates overeducation into genuine and apparent overeducation: graduates in a sub-graduate occupation who are satisfied are defined as apparently overeducated, whereas those who are dissatisfied are called genuinely overeducated. As such, job satisfaction data are needed to define the overeducation. However, only the latest wave of the survey (IFLS5 2014) that asks a question related to job satisfaction[42] in the data. Most respondents also assert that they are relatively satisfied with their job, making the respondents of those genuinely undereducated and overeducated relatively small, with only 304 and 345 respondents (out of 8489 respondents/total sample), respectively, or less than 5 per cent of the total sample in 2014. Thus, the present study only focuses on the combined general measure rather than distinguishing between the genuine and apparent overeducation.

The present study also does not consider skill mismatch, due to the unavailability of skill data. IFLS have a question on the highest education achieved, but the surveys do not have questions related to the type of education. Subsequently, horizontal mismatches cannot be explored in the present study. In addition, IFLS questionnaire has several questions related to job characteristics, such as physical conditions, skill in dealing with people, and computer capabilities, but there are no questions that are commonly used to measure skill mismatch or any direct comparisons between skill endowments and skill use[43].

---

[40] Standard deviation also has some criticisms, especially because of the arbitrary nature of the choice of cut-off points; also, if a particular occupation contains a large proportion of overeducated workers, this will raise the occupational average and the corresponding cut-off point thus underestimating the true level of the overeducation. The assumption of symmetry that the standard deviation method implies also seems unrealistic (Hartog, 1997).

[41] Mean refers to average value or sum of all of the given data; then, divided by the number of data entry or observation. While, mode refers to the number that occurs most often in the category.

[42] IFLS5 of 2014 has job satisfaction data (tk16c1): "How satisfied are you with your current job?" The alternatives for the answer are "very satisfied", "satisfied", "unsatisfied" and "very unsatisfied". Thus, "very satisfied" and "satisfied" answers can be categorised as apparently over-educated, and "unsatisfied" as well as "very unsatisfied" are grouped into genuinely mismatched. According to the results, it seems that most of the sample is satisfied with the job, with more than 80 per cent of the sample is apparently mismatched, for both undereducation and overeducation.

[43] Jones and Sloane (2010) study the disability and skill mismatch in the UK, using the 2004 British Workplace Employment Relations Survey (WERS). To measure skill mismatch, they asked questions like "How well do the work skills you personally have match the skills you need to do your present job?"

Mismatch Determinants (Explanatory Variables)

Explanatory variables here are determined by the explanatory variables mostly used in previous empirical studies (Appendix XI, Table XI.1 and Table XI.2), or the other variables relevant to the Indonesia context. Explanatory variables in this study are then classified into four categories: personal characteristics, household characteristics, work related and firm size, and area/residence variables.

Firstly, the personal characteristic variables consist of:

(1) sex (1=female),the hypotheses is: male faces a slightly higher educational mismatch risk than female, as asserted by European Commission (2012);

(2) marital status (single, married and cohabitate) with the hypothesis is married workers face a slightly higher probability of being overeducated (educational mismatch);

(3) ethnicity (1=Javanese) as majorities, the hypothesis is mismatch incidence has a greater probability for minorities; and

(4) age and the square of age, the hypothesis here is that age and overeducation have a negative relationship (Boll *et al.*, 2016). In additions, the present study prefers to choose age and square of age, rather than experience and square of experience to avoid multicollinearity problem[44].

Another potential variable is the subject of study, as explained in the overeducation determinant part, but the data are not available in all waves of the IFLS.

With regards to household characteristics, the present study chooses to use the presence of young children (0-5 years old) in the household as a variable. The hypothesis is the number of children (young children aged between 0 – 5 year old) has a significant effect on the risk of overeducation for female workers in particular.

Turning to work related and firm size, this category consists of[45]:

---

[44] Multicollinearity refers to a situation in which two or more explanatory variables in a multiple regression model are highly linearly related. The present study follows Dong's (2016) to calculate potential experience, as explained in Chapter 3, which depends on age and year of schooling. On the other hand, a match is also defined by year of schooling and its average. Choosing age (rather than potential experience) could then minimise multicollinearity problem.

[45] Another alternative variable is contract, as Boll *et al.* (2016), Green and McIntosh (2007) and Ortiz (2010), and around 30 per cent of the studies in Appendix XI consider job status. In the present study, there is question TK25A5 from IFLS5 (2014) which asks, "Do you work with a contract?" with the alternative answers being "Yes, with no fixed contract", "Yes, with fixed contract", "No" and "Do not know". However, the data from IFLS3 (2000) do not have a similar question. Thus, the present study chooses not to use contract as a variable, considering the comparability model between 2000 and 2014. Furthermore, the present study does not employ any occupation categories because the matched variable is calculated by

(1) tenure and tenure squared (as well as experience and experience squared), the hypothesis is: longer tenure reduces the risk of overeducation;
(2) working status: part time and full-time workers[46], the hypotheses are: being a part-time worker reduces the probability of being undereducated and increases the probability of being overeducated (Sloane *et al.*, 1999);
(3) sector (private/public), the hypothesis is working in the public sector could increase the likelihood of being overeducated (Ortiz, 2007);
(4) industries, educational mismatch tends to be associated with the low education level of production and agricultural workers (Allen, 2016);
(5) firm size[47], the hypothesis is workers in small firm face a higher probability being educational mismatch (Dolton and Vignoles, 2000).

The last category is region or area, which consists of urban or rural areas and provinces. Indonesia has 34 provinces which have diverse characteristics, with a wide range from medium to very high human capital development. To simplify the analysis, the present study defines two dummy variables for province: capital province (DKI Jakarta) and non-capital province (the other provinces). The capital province is considered since most companies' headquarters and central government offices are located here. Moreover, most new jobs are generated in urban areas in Indonesia (Allen, 2016). This is also based on the Harris Todaro model that rural workers have an incentive to migrate to the urban areas despite of urban unemployment, because of the potential of higher earnings in the urban area. Then, the hypothesis is: workers in urban or the capital province (which has a high unemployment rate) have a higher risk of being overeducated, while individuals living in regions with low unemployment rates are less likely to be overeducated (as per Clark *et al.* (2012) finding).

---

the difference between education qualifications that is achieved and the average education of each occupation category. This is to avoid multicollinearity problem. In fact, most of the literatures in Appendix XI use occupation as part of cohort analysis (Dolton and Silles, 2002; and Morano, 2014).

[46] full-time workers are defined as those working equal to or more than 30 hours per week, following ILO's definition.

[47] Consists of: firm with 1-19 workers (small firms), 20-99 workers (medium firms), and more than 100 workers (large firms).

4.3.3 Data

IFLS3 (2000) and IFLS5 (2014) are selected for the present study; these surveys were fielded in 2000 and in late 2014 - early 2015, respectively. The consideration in choosing these years is similar to Chapter 3; the year 2000 represents the circumstances before the Education System Law (No. 20 of 2003) was enforced and 2014-2015 represents the situations after the law had been effective; or in other words, the situations before and after the education expansion occurred in Indonesia.

Education Variable

According to previous empirical studies (Appendix XI), there are two alternatives to determine a match, *i.e.* years of education and education level. Both of these are similar, except that years of education divide university level into diploma, undergraduate, master and PhD; thus, having more detailed education segregation compared to education level. The present study prefers years of education (the upper band or highest years of each category) in line with the questions in the IFLS surveys; the highest education attained. Years of education are a continuous variable and is divided into 7 categories:

1. 0-6 years of primary education or less, represented by 6 years in the data;
2. 7-9 years for junior high school, represented by 9 years;
3. 10-12 years for senior high school, represented by 12 years;
4. 13-15 years for diploma degree, represented by 15 years;
5. 13-16 years for undergraduate degree, represented by 16 years;
6. 17-18 years for master's degree, represented by 18 years; and
7. 19-22 years for doctoral degree, represented by 22 years.

Occupation Classification

Turning to occupation classification, the IFLS has 11 categories and 86 sub-categories, with these sub-categories further defined by appropriate and adjacent occupations. The 11 categories (1-digit code) are similar to the ISCO-08 major occupation group from ILO, which are commonly used by previous studies such as: Allen (2016), Chua and Chun (2016), Lee *et al.* (2016) and Flisi *et al.* (2017).

Each category has a varied number of sub-categories, for example: the first category is the professional or category 0 which has 9 sub-categories, consisting of physical scientists, architects, surveyors, aircraft officers, medical practitioners, statisticians, and economists. The third category only has 6 sub-categories: legislative officials; managers; administrators of the government, non-government and unknown; and school principals and other managers. Meanwhile, the last category only consists of the military and police personnel. The complete sample distribution based on categories is indicated in Table 4.3 while the more detailed list of sub-categories (86 sub-categories) is shown in Appendix XII.

Using the very detailed 86 sub-categories gives a very small observation. For instance, physical scientists and related technicians only allows 2 observations[48]. On the other hand, using 11 categories would have a very wide range of means of education, subsequently mean and mode would have a significant difference and the result could not be robust. For instance, in professionals or category 0, the education mean of aircraft and ship's officers is 15.67 years; whereas the education mean of physical scientists and related technicians is substantially lower (11 years), see Appendix XII for more detailed data.

As such, the present study arranges the occupation category by collapsing some sub-categories (with low observation) which could reduce low number of observations, but at the cost of increasing heterogeneity within an occupation. Also, keeping some categories which have large number of observations will retain the homogeneity within those groups. As Foxton (2016) asserts, using an appropriate digit level of occupation category can provide a good balance between strong sample size and reducing the level of heterogeneity in roles within occupational grouping. Similarly, Nordin *et al.* (2010) point out that detailed education and occupation classifications make it possible for researchers to objectively decide whether there is a match or mismatch.

Principally, the present study keeps the sub-categories which have relatively large sample and merges the remaining (particularly sub-categories with sample size less than 30), for

[48] Regarding the sample size and the minimum sample for each category, Agresti and Min (2002) suggest a sample size of 30 as a lower bound for large-sample inference about the mean of a quantitative variable. The motivation for this size is the similarity when $n \geq 30$ between the standard normal distribution and t distribution, with a normal population. However, that simple guideline cannot apply well to all cases and may fail for high non-normal populations. Meanwhile, other studies assert that the most common is the lower bound of 10 for each category (Moore and McCabe, 1998), and 5 observations (Triola, 2000). Battu and Sloane (2002) use 2-digits classification of occupation, where occupations with fewer than 10 observations merge with the other appropriate and adjacent occupation.

example: professionals or category 0 has 151 observations in 2000 and 328 observations in 2014. The present study then re-categorises it into 3 sub-categories (sub-category 3, 7 and the remaining observations in professionals or category 0). The present study also keeps surveyors (subcategory 3) and nurses, midwives, x-ray technicians, and traditional medicine practitioners (subcategory 7) since they have more than 30 observations; and merges the remaining, which makes the total remaining of 42 and 53 observations in 2000 and 2014, respectively. From 11 occupation categories (consisting of 86 sub-categories), the present study then disaggregates them into 44 categories. The comparison list of occupations that merge and the same as the original sub-category is provided in Appendix XII Table XII-3. While, the summary statistic for this hybrid group of occupation is shown in Table 4.3.

This hybrid category indeed affects mean and/or mode within occupation group. For instance, teacher (code 13) as a part of the other professional category, has mean of years of schooling of 14.25 and mode of 12 in the year 2000. The other professional category itself previously has lower mean compared to the teacher category (12.38 year). Meanwhile, the mean of the other professionals (including teachers) now become 13.68 years. In addition, teacher has the large number of observations consisting of 451 cases, more than 75 per cent of the sample in category 1X (the other professionals). Thus, the hybrid category is better than the aggregate (1-digit category), in terms of retaining the homogeneity of sample with large observation and maintaining minimum number of observations for the category with small observation at once.

Table 4.3: Sample Distribution Based on Group of Occupation and Years of Schooling, 2000 and 2014

Occupation Code and Title		2000						2014					
		N	Mean	Mode	SD	Min	Max	N	Mean	Mode	SD	Min	Max
0X	Professional/Technical workers	42	13.31	12	3.04	6	18	53	13.92	12	2.59	6	18
03	Surveyors, draftsmen, engineering assistants	40	11.5	12	2.48	6	16	59	12.49	12	2.6	6	18
07	Nurses, midwives, x-ray technicians, traditional medicine practitioners	69	12.13	12	2.31	6	16	216	14.21	15	2.27	6	22
1X	The other professionals	117	12.38	12	2.93	6	16	123	13.31	16	3.28	6	18
13	Teacher	451	14.25	12	2.1	6	18	878	15.68	16	1.54	6	22
2X	Administrative /managerial workers	25	14.08	12	2.16	12	18	57	13.61	16	3.1	6	18
3X	Clerical and related workers	89	11.8	12	2.63	6	16	144	11.92	12	2.86	6	18
30	Clerical supervisors	78	11.96	12	3.01	6	16	70	13.26	12	2.69	6	18
31	Government executive officials	126	12.66	12	2.56	6	16	199	14.09	16	2.75	6	18
33	Bookkeepers, cashiers, and related workers	122	12.51	12	2.47	6	18	305	12.87	12	2.52	6	18
39	Clerical and related workers not elsewhere classified	160	12.59	12	2.49	6	16	354	12.62	12	2.77	6	18
4X	Sales workers	47	12.23	12	2.83	6	16	151	12.95	12	2.79	6	22
44	Insurance, real estate, securities and business services, salesman and auctioneers	69	13.51	12	2.32	6	18	147	13.97	12	2.11	9	18
45	Salesmen, shop assistants and related workers	459	9.72	12	2.85	6	16	624	10.89	12	2.85	6	18
5X	Service workers	59	8.61	6	2.8	6	16	214	10.07	12	2.94	6	16
51	Working proprietors (catering and lodging services)	167	9.84	12	2.98	6	16	266	11.48	12	2.7	6	16
54	Maids and related housekeeping service workers NEC	451	8.58	6	3.03	6	16	424	10.47	6	3.72	6	22
55	Building caretakers, char workers, cleaners and related workers	75	9.08	6	2.86	6	16	158	10.54	12	2.77	6	18
58	Protective service workers	121	10.23	12	2.56	6	16	216	11.46	12	2.3	6	16
59	Service workers not elsewhere classified	167	10.44	12	3.08	6	16	238	12.04	12	2.85	6	18
6X	Agriculture, animal husbandry, forestry workers, fishermen and hunters	141	7.72	6	2.57	6	18	154	9.82	6	3.58	6	18
62	Agricultural and animal husbandry workers	654	6.98	6	2.06	6	16	462	8.8	6	3.05	6	18
63	Forestry workers	175	8.31	6	2.72	6	16	99	9.25	6	2.79	6	18
7X	Craft and related trade workers	44	7.84	6	2.61	6	12	83	9.4	6	2.9	6	16
70	Production supervisors and general foreman	77	11.6	12	2.52	6	16	100	11.9	12	2.78	6	18

71	Miners, quarrymen, well drillers and related workers	42	8.6	6	2.84	6	16	36	10.22	6	3.83	6	18
74	Chemical processors and related workers	43	8.95	6	3.13	6	16	65	9.94	12	2.88	6	18
75	Spinners, weavers, knitters, dyers, and related workers	115	8.57	6	2.67	6	16	80	9.31	12	2.88	6	16
77	Food and beverage processors	232	9.15	6	3.17	6	16	328	10.43	12	3.1	6	22
79	Tailors, dressmakers, sewer, upholsterers and related workers	189	8.5	6	2.68	6	16	257	9.44	12	2.49	6	16
8X	Plant and machine operators and assemblers	66	9.05	12	2.71	6	12	97	10.92	12	2.65	6	16
80	Shoemakers and leather goods makers	34	8.82	6	2.66	6	15	59	10.34	12	2.54	6	16
81	Cabinet makers and related wood makers	99	8.46	6	2.65	6	16	92	10.53	12	2.84	6	16
84	Machinery fitters, assemblers, repairers and precision instrument makers (except electrical)	78	10.09	12	2.58	6	16	164	11.12	12	2.31	6	16
85	Electrical fitters and related electrical and electronics workers	42	11.05	12	2.54	6	16	60	10.78	12	2.73	6	18
87	Plumbers, welders, sheet-metal and structural metal preparers and erectors	32	8.53	6	2.65	6	12	42	9.55	12	2.91	6	16
9X	Elementary occupation	60	8.6	9	2.24	6	12	95	9.67	12	2.71	6	16
92	Printers and related workers	73	9.71	12	3.03	6	16	101	10.41	12	2.76	6	16
94	Production and related workers not elsewhere classified	48	8.31	6	2.64	6	15	62	9.69	6	3.11	6	16
95	Bricklayers, carpenters and other construction workers	519	7.99	6	2.69	6	16	425	9.49	6	3.06	6	18
97	Material handling and related equipment, operators, dockers and freight handlers	137	8.86	6	3.14	6	16	249	10.39	12	3.19	6	18
98	Transport equipment operators	339	8.81	6	2.66	6	16	304	10.19	12	2.77	6	18
99	Labourers not elsewhere classified	155	9.13	6	3.25	6	16	158	10.68	6	3.9	6	22
M	Armed forces occupation	57	11.18	12	2.21	6	18	21	13.29	12	1.87	12	16
		6385						8489					

Source: The author’s calculation.

Note: Code number and X (*e.g.* 0X and 1X) refers to occupation category. 2 digits code (*e.g.* 03 and 62) refers to occupation sub-category. N is the number of observations, SD: standard deviation, min: minimum value of each occupation category, max: maximum value of each category. Occupation code is based on 1 and 2 digits of IFLS classification. The calculation of mean, mode and other descriptive statistics presented in this table are for the years of schooling variable.

Sample Restrictions

Sample restrictions are also applied for the estimations to ensure that the individuals are in the labour market (Table 4.4). The first restriction is age; between 16-55 years old. This restriction considers the age of finishing compulsory education and the retirement or state pension age, similar to previous studies (*e.g.* Dockery and Miller, 2012). The actual sample of IFLS range from 14 years of age or 14+.

The next restriction is employment status (non-missing data); it takes away around 17 per cent of the sample in 2014, and even bigger in 2000 (30 per cent). The third restriction is wages, *i.e.* the waged sector (both public and private) considering the same characteristics (receiving regular wages). Here, 31 per cent and 54 per cent of the sample are lost in 2014 and 2000 respectively. Other possible reasons are because most of the individuals are casual workers (not included in the sample from 2000), the proportion of self-employment is relatively high, or the proportion of sample which does not fill their employment status is relatively large. The remaining (non-missing personal characteristics/HH/job/area) is not absolute restriction. Some losses also occur because of missing data, though it is not really significant compared to the previous restrictions.

Table 4.4: Sample Restrictions

	2000	**% Lost**	**Waged and Casual**		**Waged Only**	
			2014	**% Lost**	**2014**	**% Lost**
All sample	25,825		36,381		36,381	
Age 16-55	21,100	-18.30	29,797	-18.10	29,797	-18.10
Employment status data (non-missing)	14,771	-30.00	24,473	-17.87	24,473	-17.87
Employment status (Government workers, Private workers, Casual workers not in agriculture and in agriculture) ==> without self-employment			12,698	-48.11	12,698	-48.11
Employment status (Government workers, Private workers)	6,780	-54.10			8,712	-31.39
Mismatch (non-missing) and Personal characteristics	6456	-4.78	10,852	-14.54	8,712	0.00
Mismatch (non-missing) and HH (Number of children)	6456	0.00	10,852	0.00	8,712	0.00
Mismatch (non-missing) and Job characteristics	6385	-1.10	10,594	-2.38	8,498	-2.46
Mismatch (non-missing) and Residence/Area	6385	0.00	10,594	0.00	8,498	0.00

Source: The author’s calculation.

Furthermore, waged sector and casual data in 2014 are used to test the sensitivity of the results. From 25,825 individuals in 2000 and 36,385 individuals in 2014, this study has the final sample of 6,385 individuals and 8,488 individuals in 2000 and 2014, respectively; or around 23-25 per cent of the total sample in each year data (Table 4.4).

Turning to casual workers, it is defined as workers who do not have regular or systematic hours of work or an expectation of continuing the work. A typical casual employee is employed on a daily basis when the employer's need arises. Statistics Indonesia defines casual workers as those who have no permanent employers for a certain time, for example less than three months for the construction sector. Casual workers generally have an education attainment of junior high school or less and receive comparatively low levels of remuneration.

Casual workers and unpaid workers are also associated with informal employment. Informal employment can be classified into informal self-employment and informal wage-employment. Under informal self-employment are employers in informal enterprises, own-account workers in informal enterprises, unpaid family workers and members of informal producers' cooperatives. Informal wage-employment includes employees who are employed either in formal or informal enterprises without formal contracts, worker benefits, or social protections. In addition, informal enterprises are household enterprises engaged in the production of goods or services with the primary objective of generating employment and incomes for the persons concerned, and which typically operate at a low level of organisation, with little or no division between labour and capital as factors of production, and on a small scale. Labour relations—where they exist—are based mostly on casual employment, kinship, or personal and social relations rather than contractual arrangements with formal guarantees (Cuevas, *et. al.*, 2009). Allen (2016) also adds that one of the industries that have a large proportion of casual workers, particularly in the non-agriculture industry, is the construction industry.

According to ILO (2002), informal sector can have different definition across the countries, as indicated in the latest ILO published data on informal sector enterprises that are based on information from 54 countries; most of which still adhere to their own national definitions of the informal sector which are not entirely in line with the international statistical definition adopted by the 15th ICLS and the 1993 System of National Accounts (1993 SNA). Under the 1993 definition, only one category of informal wage workers is counted, namely employees of informal sector enterprises, and individual countries can decide what size of unregistered units to include in the informal sector and whether the agricultural sector and domestic workers should be covered. Rothenberg *et al,* (2016) refer informal sector as all economic activities conducted by firms that are not formally registered with the government and do not pay taxes. Statistics Indonesia defines casual workers that do not have permanent employers for a certain time. ADB (2009) emphasise informal employment is bound more by social relations rather than contractual arrangements with formal guarantees. While, Cuevas *et al,* (2009) argue given the limitation of the data, informal employment in Indonesia is defined as employment on casual basis and unpaid work.

Considering IFLS data, an individual in the survey could choose between several employment sectors if he/she was determined as working (TK24A):

1. self-employed (without help),
2. self-employed with the help of householders/temporary workers,
3. self-employed with the help of regular workers,
4. government worker/employee,
5. private worker/employee,
6. unpaid family worker.

In the 1997 survey, all the self-employed sectors are combined into one sector called the self-employed. However, IFLS adds two additional sectors in 2007:

7. casual worker in agriculture, and
8. casual worker in non-agriculture.

Casual workers in the present study refer to category 7 and 8 (casual workers in agriculture and non-agriculture). However, the IFLS surveys do not provide any specific definitions on these employments. Thus, the present study follows Cuevas *et al,* (2009), informal sector covers unpaid family workers and casual workers both in agriculture and non-agriculture or categories 6, 7 and 8. In 2014/15, IFLS5 has a total sample of 36,381 individuals, 27.43 per cent of the sample is waged workers, 25.27 per cent is self-employed, 7.12 per cent is unpaid family worker, and 7.47 per cent is casual worker both in agriculture and non-agriculture; Thus, the proportion of informal sector in 2014 is around 14.6 per cent.

4.3.4 Summary Statistics

This part will focus on the dependent variables for all individuals, based on gender and sector. But before getting there, the comparison between mode and mean, as well as the normality test will be discussed. The present study chooses mode as the main model or a better measurement of education mismatch than mean because mode is more suitable for categorical data like the match variable. Moreover, mode is a better measure than mean when the data are skewed. Table 4.5 shows the distribution and normality test of the match variable. The standard deviation and variance of mode of the model are slightly higher than the mean of the model for each period. Meanwhile, skewness[49] and kurtosis[50] of mode model are slightly lower than mean model. The skewness of mode is a negative value, implying that the distribution is skewed to the left. This is in contrast to mean model; the skewness is positive, or that the distribution is skewed to the right. According to the rule of thumb of normal distribution (skewness value is 0 and kurtosis is 3), only the mean model of 2014 is eligible. The present study also conducts Skewness/Kurtosis tests for Normality in Stata, and the result confirms that only the mean model of 2014 that has a normal distribution; the mode model of both years and the mode of 2014 are not normally distributed. Thus, mode method is better for this case (not normally distributed), as is in line with Lund Research (2013) that mode is a better measure than mean when the data are skewed. In addition, in symmetric unimodal distributions (such as the normal distribution), the mean, median and mode all coincide (Lane, 2007). Thus, mode model is advantageous for both normally and non-normally distributed data.

The mode level of education could be a better measurement than the mean level of education in measuring required education, since it captures the most common level of education in a particular occupation, with the assumption is no bimodality in any occupation. This is in line with Kiker *et al.* (1997). Another advantage of using mode is that it is less sensitive to outliers and to technological and workplace changes (Battu and Sloane, 2002).

---

[49] Skewness is a measure of the symmetry in a distribution; a symmetrical data (normal distribution) will have a Skewness of equal to zero.

[50] Kurtosis is a measure of the combined sizes of the two tails. It measures the amount of probability in the tails. The value of kurtosis in a normal distribution is equal to 3.

Table 4.5: Normality Test of Match Variable

	2000		2014	
	Match (Mode)	Match (Mean)	Match (Mode)	Match (Mean)
Obs	6385	6385	8489	8489
Mode/Mean	2.15	2.04	2.01	1.99
SD	0.63	0.55	0.68	0.57
Variance	0.40	0.30	0.46	0.32
Skewness	-0.13	0.03	-0.01	0.00
Kurtosis	2.43	3.34	2.17	3.13
Skewness/Kurtosis tests for Normality				
Pr(Skewness)	0.00	0.40	0.78	0.96
Pr(Kurtosis)	0.00	0.00	0.00	0.02
Joint				
Prob>chi2	0.00	0.00	0.00	0.07
	not normally distributed	not normally distributed	not normally distributed	normally distributed

Source: The author's calculation.

This part answers the first, second and third research questions all at once; firstly, does education mismatch (both undereducation and overeducation) exist in the waged sector in Indonesia? Secondly, what are the estimated proportions of education mismatch in 2000 and 2014 periods? And finally, how does education mismatch change between these periods?

Based on the mismatch definition (mode model)[51], around 53-58 per cent of the sample was in the match category for both periods. In 2000, the proportion of undereducation was 13.6 per cent which then increased to 22.8 per cent in 2014. This implies that there was a significant increase in undereducation during the periods. In contrast, overeducation proportion in 2000 was around 28.5 per cent, which then decreased to 23.4 per cent in 2014 (Table 4.6).

This finding confirms that education mismatch indeed occurred in Indonesia in 2000 and 2014. Overall, education mismatch slightly increases during these periods (from 42 per cent to 46 per cent), and this increase is driven by the increase in undereducation. The finding also agrees with ILO's (2017) finding; an increasing trend in undereducation and a decreasing one in overeducation. This finding is also in line with the technological

[51]Individuals are considered undereducated (category 1) if they have years of education (education level) more than one standard deviation below their occupation's mode; and overeducated (category 3) if they have years of education (education level) more than one standard deviation above their occupation's mode (adapted from Dockery and Miller, 2012).

change theory that while technology changes rapidly, better-educated workers cannot be made instantaneously - the mean of years of schooling was 9.7 years in 2000, which then increased to 11.5 years in 2014. In other words, the increase amounts to only around 2 years of schooling during the periods. In addition, education quality in Indonesia is relatively low compared to other countries in the region. Meanwhile, McKinsey (2012) estimates that the demand for semi-skilled and skilled workers (reflected by high education qualifications) may rise to 113 million by 2030. As a result, undereducation incidence would arise.

Arguably, this trend could also be due to underproduction and high demand for workers with higher education qualifications, characterised by high wage premium (Carnevale and Rose, 2013). Although the trend of return to education decreases between 2000 and 2014, workers with university qualifications still obtain relatively higher wage premium compared to workers with lower qualifications (see Chapter 3). Moreover, workers with junior high school or lower education are likely to shift between working as production workers and agriculture labours throughout the year, but these workers find it more difficult to climb the career ladder.

Besides, the minimum wage policy could also affect undereducation. The minimum wage in Indonesia is determined at province, district and occupational levels (Article 89 of Labour law Number 13/2003), as explained in Chapter 2. As a result, if workers with primary school and senior high school qualifications apply for the same occupation (*e.g.* blue-collar works in textile factories), they will be offered the same amount of wages, which is equal to the minimum wage. As such, this policy could dishearten individuals to pursue higher education level and increases undereducation coincidence.

The mixed effect of the increasing minimum wage in Indonesia is also found by Smeru (2002). Smeru finds that the increase in minimum wages push up wages of blue-collar workers. Moreover, there is a positive relationship between minimum wages and average wages of most other groups of workers, such as females, the youth, the less educated and white-collar workers. In contrast, the increase in minimum wages have a negative impact on urban and the waged sector employment, with the exception of white-collar workers. Also, Del Carpio *et al.* (2015) find that the effect of minimum wages in Indonesia are more binding in small firms. Though their study does not provide an analysis based on education level, it shows that the effect of minimum wages could be different among groups of workers.

And finally, another argument is related to the method used in this research: the weakness of the Realise Method (RM) is that it only considers the supply side, so the changes in the mode only reflects the changes in average workers' education. The measurement cannot represent the changes because the jobs demanded higher education qualifications. Also, changes in education levels in the economy imply that employers will allocate workers differently (Mason, 1996).

By using the mean, the estimation provides a relatively higher proportion of matched category than the estimation by mode; the match of mean model is more than 68 per cent, meanwhile the match in the mode model was around 53 per cent in 2014; a similar pattern also occurred previously in 2000. Overeducation and undereducation proportions are relatively lower when estimated by the mean model than by the mode model. Nevertheless, the conclusion for both mode and mean is the same; there was an increase in undereducation, as well as a decrease in overeducation, and a decrease in education matching in aggregate between 2000 and 2014 periods.

It seems that the decrease in overeducation and the increase in undereducation also occur when the sample is separated by gender and by sector, except for the public sector. In the public sector, there was an increase in overeducation, from 29.32 per cent to 31.81 per cent, as well as a decrease in undereducation from 15.51 per cent to 10.82 per cent between 2000 and 2014 periods. In line with the findings in Chapter 3, it is possible that the public sector prefers workers with high education levels: those with at least senior high school qualifications. For example, teacher's education requirement is at least Diploma II for rural areas and Diploma IV or undergraduate in general. For unskilled jobs such as cleaning, the public sector prefers outsourcing or using private firms' services, rather than hiring directly. The present study will elaborate this further in the part about estimation by sector, but this finding could also imply that the private sector is the main contributor in the increase of undereducation.

In terms of gender, the present study confirms the ILO's finding that male workers tend to have lower education levels than female workers, in particular for education of at least 13 years or university level (see Appendix XIII). In terms of the trend of mismatches, the increase in undereducation was around 38-42 per cent for both males and females. Males experienced a substantial decrease in overeducation from 31.8 per cent to 25.5 per cent while overeducated female decreased from 22 per cent to 20 per cent (Table 4.6). In other words, the pattern persists throughout the sample; a decrease in overeducation and an increase in undereducation.

Table 4.6: Match Based on Mode and Mean (in per cent)

	Match (Mode)							
	2000				2014			
Year	UE	M	OE	Total	UE	M	OE	Total
All Individuals								
Number of obs.	869	3,697	1,819	6,385	1,931	4,576	1,982	8,489
Proportion	13.61	57.90	28.49	100.00	22.75	53.91	23.35	100.00
Males								
Number of obs.	564	2,301	1,339	4,204	1,127	2,712	1,316	5,155
Proportion	13.42	54.73	31.85	100.00	21.86	52.61	25.53	100.00
Females								
Number of obs.	305	1,396	480	2,181	804	1,864	666	3,334
Proportion	13.98	64.01	22.01	100.00	24.12	55.91	19.98	100.00
Public								
Number of obs.	156	555	295	1,006	152	806	447	1,405
Proportion	15.51	55.17	29.32	100.00	10.82	57.37	31.81	100.00
Private								
Number of obs.	713	3,142	1,524	5,379	1,779	3,770	1,535	7,084
Proportion	13.26	58.41	28.33	100.00	25.11	53.22	21.67	100.00

	Match (Mean)							
	2000				2014			
Year	UE	M	OE	Total	UE	M	OE	Total
All Individuals								
Number of obs.	828	4,478	1079	6,385	1,384	5,774	1,331	8,489
Proportion	12.97	70.13	16.90	100.00	16.30	68.02	15.68	100.00
Males								
Number of obs.	567	2,854	783	4,204	780	3,491	884	5,155
Proportion	13.49	67.89	18.63	100.00	15.13	67.72	17.15	100.00
Females								
Number of obs.	305	1,396	480	2,181	604	2,283	447	3,334
Proportion	13.98	64.01	22.01	100.00	18.12	68.48	13.41	100.00
Public								
Number of obs.	154	621	231	1,006	64	952	389	1,405
Proportion	15.31	61.73	22.96	100.00	4.56	67.76	27.69	100.00
Private								
Number of obs.	674	3857	848	5,379	1,320	4,822	942	7,084
Proportion	12.53	71.70	15.77	100.00	18.63	68.07	13.30	100.00

Source: The author's calculation.
Note: UE: undereducated; M: matched; OE: overeducated.

In terms of changes, the present study performs t-test to ensure that there is an equality of means between the data from 2000 and 2014. The hypothesis (H0) is that the difference between mean of 2014 and mean of 2000 is equal to zero. The t-test is conducted in STATA and the result indicates that H0 is rejected, or the difference of means in the mismatch between 2014 and 2000 are different from zero (all is significant at 5 per cent). Thus, this result implies that the increase in the fraction of undereducation is statistically significant. Likewise, the decrease in the fraction of overeducation is also statistically significant. The t-test result for all individuals is shown in Table 4.7 and the remaining results are provided in Appendix XIV[52].

Although the mean and mode results are slightly different, the present study finds a similar pattern of education mismatch; an increase of undereducation and a decrease of overeducation, except for the public sector. Hence, the mean or mode methods give the same conclusions of education mismatch in Indonesia between 2000 and 2014 periods[53].

Table 4.7: T-Test of the Main Model (All Individuals)

	Mean		Difference	T-Statistic	Degree of Freedom	P-value
	2000	2014				
Main Model (Mode)						
UE	0.14	0.23	0.09	14.20	14872	***
M	0.58	0.54	-0.04	-4.86	14872	***
OE	0.28	0.23	-0.05	-7.13	14872	***

Source: The author's calculation.

Table 4.8 shows the summary statistics of all individuals from both periods. The dependent variable of the main model is match (mode), and the number of observations in 2000 and 2014 were 6,385 and 8,489 - respectively. The alternative dependent variable or match (mean) will be discussed further in the sensitivity test part. Turning to the independent variables, this part highlights some findings based on sex, age, tenure, urban and capital (residence in capital city). Most of the observations are males; females were only 34 per cent in 2000 and 39 per cent in 2014. As age is the first restriction in the present study, the minimum age was 16 and the maximum was 55. In 2000, the average

52 The other results are similar to the main model result, except for the increase in fraction of overeducation in public sector which seems to be insignificant.

53 The present study also estimates the mismatch based on 1-digit occupation classification (the aggregate one) and finds an inconsistent result between mode and mean (see Appendix XII).

age of the waged sector was 31.66 years, increased to 33.28 years in 2014. Meanwhile, the tenure variable had a minimum value of 0 year and a maximum value of 50 and 40 years for 2000 and 2014, respectively. The maximum age was 55 years while the maximum tenure was 50 years, which is nonsense. However, only 4 out of 6,385 individuals in 2000 (less than 1 per cent) had tenure of more than 40 years. As such, those negative values were set to zero. The average tenure in 2000 was 6.56 years, which slightly decreased to 6.34 in 2014. In terms of sector, more than 80 per cent of the sample in both periods were from the private sector. Most of the sample also lived in urban areas; more than 60 per cent of the sample which is slightly higher than the data from Statistics Indonesia (around 55 per cent living in urban areas). Moreover, only around 10 per cent of the sample lived in the capital city.

Table 4.8: Summary Statistics, All Individuals

Variable	2000				2014			
	Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Dependent Variable:								
Match (mode)	2.15	0.63	1	3	2.01	0.65	1	3
Alternative Dependent Variable:								
Match (mean): Sensitivity Test	2.04	0.55	1	3	1.99	0.57	1	3
Independent Variables:								
Sex (1=Female)	0.34	0.47	0	1	0.39	0.49	0	1
Marital Status	1.74	0.52	1	3	1.82	0.48	1	3
Ethnicity (1=Javanese)	0.62	0.49	0	1	0.45	0.5	0	1
Age	31.66	9.91	16	55	33.28	9.57	16	55
Age squared	1100.5	677.94	256	3025	1198.89	676.48	256	3025
Young children (0-5 years old)	0.14	0.38	0	2	0.36	0.58	0	3
Tenure	6.56	7.45	0	50	6.34	7.22	0	40
Tenure squared	98.53	198.88	0	2500	92.36	192.04	0	1600
Status: part time	0.15	0.35	0	1	0.17	0.38	0	1
Sector: private	0.84	0.36	0	1	0.83	0.37	0	1
Industry	5.54	3.01	1	9	6.18	2.78	1	9
Firm Size	1.4	0.65	1	3	1.68	0.8	1	3
Urban	0.63	0.48	0	1	0.72	0.45	0	1
Capital	0.13	0.34	0	1	0.09	0.28	0	1
Number of Obs.	6385				8489			

Source: The author's calculation.

The present study also extends the analysis based on sub-sector (Table 4.9) and finds that there is no substantial difference between the genders in terms of the summary statistics. Thus, it is not surprising if gender does not have a significant effect on education mismatch and its determinants. In terms of sector, the present study highlights at least two variables having a significant difference between the private and public sectors, *i.e.* age and tenure. The public sector had a significantly higher age average (38-39 years) than the private sector (30-32 years). Generally, the private sector employs younger workers than the public sector. Similarly, the public sector has a significantly longer tenure than the private sector. The average tenure in the public sector was around 12 years while the private sector had tenure of around 5-6 years between those periods. However, the longer tenure and older age do not necessarily suggest a lower mismatch in the public sector; the present study will elaborate this issue further in the sector analysis.

Table 4.9: Summary Statistics (Mode Model) based on Sub-categories, 2000 and 2014

	2000		2014		2000		2014	
Variable	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
	Female				**Male**			
Match (Mode)	2.08	0.59	1.96	0.66	2.18	0.65	2.04	0.69
Sex (1=Female)	1.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
Marital Status	1.76	0.61	1.85	0.52	1.74	0.47	1.80	0.45
Ethnicity (1=Javanese)	0.64	0.48	0.47	0.50	0.61	0.49	0.44	0.50
Age	30.54	9.72	32.52	9.59	32.24	9.96	33.76	9.53
Age squared	1027.30	646.46	1149.76	671.33	1138.47	690.73	1230.67	677.95
Young children (0-5 years old)	0.13	0.36	0.31	0.55	0.14	0.38	0.40	0.59
Tenure	6.11	7.16	6.00	7.10	6.80	7.58	6.56	7.29
Tenure squared	88.53	187.25	86.40	191.45	103.72	204.49	96.21	192.34
Status: part time	0.20	0.40	0.25	0.43	0.12	0.32	0.13	0.33
Sector: private	0.84	0.36	0.81	0.39	0.84	0.37	0.85	0.36
Industry	5.68	3.08	6.58	2.71	5.47	2.96	5.93	2.80
Firm Size	1.45	0.69	1.64	0.80	1.37	0.63	1.70	0.80
Urban	0.67	0.47	0.72	0.45	0.60	0.49	0.71	0.45
Capital	0.15	0.36	0.08	0.27	0.12	0.33	0.09	0.29
Number of Obs.	2181		3334		4204		5155	
	Private				**Public**			
Match (Mode)	2.15	0.63	1.97	0.68	2.14	0.66	2.21	0.62
Sex (1=Female)	0.34	0.47	0.38	0.49	0.34	0.47	0.45	0.50
Marital Status	1.71	0.54	1.80	0.50	1.92	0.37	1.93	0.37
Ethnicity (1=Javanese)	0.64	0.48	0.48	0.50	0.52	0.50	0.32	0.47
Age	30.27	9.50	32.32	9.26	39.09	8.68	38.08	9.67
Age squared	1006.49	637.20	1130.48	639.20	1603.15	668.23	1543.84	750.12
Young children (0-5 years old)	0.14	0.37	0.35	0.57	0.17	0.40	0.41	0.62
Tenure	5.39	6.56	5.36	6.30	12.80	8.70	11.30	9.26
Tenure squared	72.14	171.62	68.35	152.72	239.60	264.59	213.40	296.19
Status: part time	0.15	0.35	0.16	0.36	0.15	0.36	0.26	0.44
Sector: private	1.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
Industry	5.11	2.91	5.72	2.72	7.89	2.38	8.54	1.68
Firm Size	1.38	0.65	1.66	0.81	1.47	0.64	1.74	0.74

Urban	0.62	0.49	0.73	0.45	0.65	0.48	0.65	0.48
Capital	0.14	0.35	0.10	0.30	0.07	0.25	0.03	0.18
Number of Obs.	5379		7084		1006		1405	

Source: The author's calculation.

4.4 Estimation Result: Education Mismatch Determinants

The main model for the analysis is mismatch based on mode (MNL). The estimations in the main model show that education mismatch in Indonesia is determined by personal and household characteristics, work-related and firm size as well as area of residency variables, all of which are observed and analysed in this research. The analysis of this chapter will now begin with undereducation, followed by overeducation as well as gender and sector analysis.

4.4.1 Undereducation

Table 4.10 presents the estimation result of Multinomial Logit Model (MNL) for all individuals based on mode in 2000 and 2014. Most variables observed in the present study significantly determine the probability of undereducation; except for ethnicity, tenure square and capital area.

Table 4.10: Education Mismatch, MNL (Mode)

Mode	2000				2014			
	Coef.	Std. Err.	P>z	ME	Coef.	Std. Err.	P>z	ME
UE								
Sex (1=Female)	-0.454	0.090	***	-0.03	0.045	0.060		0.01
Married and cohabitate	0.346	0.130	***	0.04	0.416	0.095	***	0.08
Other (Separated, divorced and widowed)	0.610	0.218	***	0.08	0.458	0.154	***	0.09
Ethnicity (1=Javanese)	-0.122	0.085		-0.01	0.005	0.057		0.00
Age	-0.121	0.034	***	0.00	-0.059	0.024	**	0.00
Age squared	0.002	0.000	***		0.001	0.000	***	
Young children (0-5 years old)	-0.020	0.110		0.00	-0.145	0.054	***	-0.02
Tenure	0.038	0.018	**	0.00	-0.017	0.012		-0.00
Tenure squared	-0.001	0.001			0.000	0.000		
Status: part time	0.217	0.115	*	0.03	-0.066	0.080		-0.00
Sector: private	-0.028	0.130		0.02	0.742	0.110	***	0.13
Industry2: mining and quarrying	2.898	0.938	***	0.03	0.747	0.224	***	0.23

Industry3: manufacturing	4.195	0.588	***	0.12	1.512	0.133	***	0.31
Industry4: electricity, gas and water	5.539	0.695	***	0.37	1.733	0.301	***	0.14
Industry5: construction	1.785	0.682	***	0.01	-0.753	0.250	***	-0.05
Industry6: wholesale, retail, restaurants and hotels	5.834	0.588	***	0.47	1.312	0.135	***	0.20
Industry7: transportation, storage, and communications	2.913	0.660	***	0.03	1.407	0.189	***	0.22
Industry8: Finance, insurance, real estate and business services	4.170	0.689	***	0.12	0.643	0.179	***	0.05
Industry9: Social services	4.271	0.585	***	0.16	0.677	0.131	***	0.08
Firm size2: 20-99 people	-0.439	0.110	***	-0.05	-0.452	0.072	***	-0.08
Firm size3: >= 100 people	-0.383	0.152	**	-0.05	-0.572	0.080	***	-0.11
Urban	-0.211	0.093	**	-0.03	-0.335	0.063	***	-0.07
Capital	-0.062	0.127		-0.01	-0.012	0.101		-0.01
Constants	-3.659	0.776	***		-1.804	0.398	***	
OE								
Sex (1=Female)	-0.403	0.070	***	-0.06	-0.120	0.061		-0.00
Married and cohabitate	-0.318	0.089	***	-0.07	-0.481	0.096	***	-0.07
Other (Separated, divorced and widowed)	-0.882	0.211	***	-0.16	-0.658	0.192	***	-0.11
Ethnicity (1=Javanese)	-0.117	0.064	*	-0.02	0.013	0.061		0.00
Age	0.144	0.026	***	0.00	0.121	0.028	***	0.00
Age squared	-0.002	0.000	***		-0.002	0.000	***	
Young children (0-5 years old)	-0.265	0.086	***	-0.05	0.026	0.054		0.01
Tenure	-0.047	0.014	***	-0.01	-0.011	0.013		-0.00
Tenure squared	0.001	0.001			-0.000	0.000		
Status: part time	-0.376	0.097	***	-0.07	-0.420	0.088	***	-0.05
Sector: private	-0.797	0.100	***	-0.15	-0.638	0.086	***	-0.13
Industry2: mining and quarrying	0.693	0.298	**	0.14	-0.184	0.253		-0.08
Industry3: manufacturing	-0.050	0.103		-0.05	-0.956	0.136	***	-0.17
Industry4: electricity, gas and water	-0.532	0.455		-0.20	-0.484	0.331		-0.10
Industry5: construction	0.185	0.120		-0.04	0.272	0.156	*	0.07
Industry6: wholesale, retail, restaurants and hotels	-1.196	0.152	***	-0.29	-0.488	0.131	***	-0.11
Industry7: transportation,	0.593	0.139	***	0.12	-0.432	0.213	**	-0.11

storage, and communications								
Industry8: Finance, insurance, real estate and business services	0.009	0.235		0.04	0.482	0.141	***	0.08
Industry9: Social services	-0.826	0.103	***	-0.19	-0.130	0.115		-0.04
Firm size2: 20-99 people	0.246	0.075	***	0.05	0.230	0.071	***	0.05
Firm size3: >= 100 people	0.471	0.104	***	0.10	0.522	0.083	***	0.10
Urban	0.549	0.072	***	0.01	0.544	0.074	***	0.08
Capital	0.342	0.093	***	0.02	-0.217	0.108	**	-0.03
Constants	-1.560	0.411	***		-2.631	0.454	***	
Number of obs.	6385				8489			
LR chi2(46)	1663.37				1357.57			
Prob > chi2	0				0			
Pseudo R-Squared	0.1378				0.0816			

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Pseudo R-Squared of 2014 is slightly lower than Pseudo R-Squared of 2000. The Pseudo R-Squared is treated as a measure of effect size, similar to how R-Squared is treated in standard multiple regressions (Starkweather and Moske, 2011). The decrease of Pseudo R-Squared in 2014 could imply that there are other unobserved variables which affect education mismatch.

The first variable is sex (female). The coefficient of this variable was negative and significant in 2000 (negative relative probability of undereducation rather than match), which indicates that females are less likely to be undereducated than males. The marginal effect was -0.03 which indicates that the probability of undereducation is, on average, about 3 percentage points lower for females than for males. There are some advantages for female workers compared to males, as World Bank (2011) documents: (1) garment companies prefer to recruit female for the sewing jobs; and (2) bank and teaching jobs are considered safe and respectable for women. Thus, firms may consider females with lower education qualifications for a job with higher education requirement to fill in some job categories. These assertions are supported by the coefficients of manufacturing and finance, insurance, as well as real estate and business services which are positive and significant, suggesting that those working in these industries are more likely to be undereducated, relative to the agriculture industries. Skill could also affect this mismatch calculation. However, the present study is limited by the lack of skill data. The analysis could have been more comprehensive otherwise. That being said, sex turned insignificant in 2014.

The married and cohabitate status, as well as the other marital status, had a positive and significant coefficient in both periods. The marginal effect of the married variable in 2000 was 4 per cent, which increased to 8 per cent in 2014. For household characteristics, the only variable observed is the number of young children aged between 0-5 years in the family. The result shows that the presence of young children in the family was insignificant in 2000, but turned negative and significant in 2014, with the marginal effect equal to -0.02. To provide a more detailed analysis, the present study will further elaborate this issue in the estimation by gender, in the gender part.

In work-related category, most variables were significant in both periods. Tenure had a marginal effect of 0.00, implying the relative probability of undereducation was the same for any job-tenure. An example is garment companies (World Bank, 2012); as long as the workers have sewing skill, tenure does not have any impacts. Unfortunately, IFLS data do not have any information on skill, thus the present study cannot explore this issue.

The result also shows that the private sector experiences a higher relative probability of undereducation incident. This confirms the finding in Table 4.6 that the increase of undereducation incident is partly driven by the private sector. The marginal effect of the private sector was 13 per cent in 2014. A possible explanation is that market mechanism occurs in the private sector. When most workers have lower education levels, companies will hire the best ones from the applicants. In addition, around 39 per cent of the sample in 2014 had 6 and 9 years of education (equal to primary and junior high school qualifications). Also, around 42 per cent of the sample had 12 years of education (senior high school qualifications). Thus, the sample is dominated by those with lower education qualifications. The present study will elaborate this issue further in the section on estimation by sector.

Moreover, most of the industry dummies are positive and significant. This suggests that mismatch (undereducation) is more likely to occur in non-agriculture sectors. This is in line with Allen (2016) who asserts that undereducation is a challenge in higher-level occupations. On the other hand, most workers with higher education qualifications work in the non-agriculture industry.

In terms of firm size, the coefficients were negative and significant, thus suggesting that those working in medium and large firms were less likely to be undereducated. Also, the marginal effect decreases with company size. For example, the marginal effect of medium firms was -8 per cent and for the large firm -11 per cent. Yin (2016) asserts that workers in large firms have more opportunities to be promoted, have better career prospects and receive more fringe benefits compared to those who work in smaller companies. Allen and Kyloh (2016) further that several large companies in the private sector have also developed their own training centres in order to meet their specialised training needs. In other words, although the workers do not have sufficient education qualifications, the firms could provide the necessary on-the-job training before they start working.

For area variables, only the urban area category that had a negative and significant coefficient in both models and in both periods, with the marginal effect changing from -3 per cent in 2000 to -7 per cent in 2014. This suggests that those working in urban areas were less likely to be undereducated compared to those working in rural areas. A possible explanation is that most new jobs are generated in the urban areas in Indonesia (Allen, 2016).

Furthermore, the Pseudo R2 of 2000 and 2014 models are 0.14 and 0.08, respectively. The interpretation is not as straightforward as the R-squared of the OLS model. The Pseudo R-squared is defined as the proportion of the variance of the latent variable that is explained by the covariate. It is basically the change in terms of log-likelihood from the intercept-only model to the current model. Nevertheless, the higher is still the better fit, but they should be interpreted with caution. It is worth noting that the Pseudo R-Squared are useful tools in evaluating multiple models predicting the same outcome on the same dataset, but they cannot be interpreted independently or compared across different datasets. In other words, a pseudo R-squared statistic without context has little meaning. A pseudo R-squared only has meaning when compared to another pseudo R-squared of the same type, on the same data, predicting the same outcome (Abdulhafedh, 2017). In the present study, the Pseudo R-Squared of 2000 and 2014 cannot be compared and interpreted independently since both of periods have different dataset.

Finally, the Likelihood Ratio chi-square test is alternative test of goodness-of-fit. The LR chi2 test shows that both equations (undereducation relative to match and overeducation relative to match) have at least one of the predictors' regression coefficients not equal to zero. Prob > chi2 is the probability of getting LR test as extreme as the observed conditions under the null hypothesis (all the regression coefficients across both models are simultaneously equal to zero). The small p value leads to the conclusion that at least one of the regression coefficients in the model is not equal to zero.

4.4.2 Overeducation

Turning to the overeducation analysis, the female variable had a negative and significant coefficient in 2000, turning insignificant in later period. This is similar to the undereducation estimation (Table 4.10). Again, this could be due to several advantages possessed by female workers, as discussed previously in the undereducation section.

The married and other marital status variable has the opposite sign of coefficient of the undereducation variable. The relative probability of overeducation was negative and significant relative to match. The marginal effect of the married status was -7 per cent in both periods. Most of the previous studies elaborate the analysis with interaction terms between marital status and sex or other variables. However, a cautious interpretation is needed here, as the data show that more than 65 per cent of the sample was married in both periods. A possible explanation is that the large number of vacancies in large labour markets is offset by a larger number of job searchers (McGoldrick and Robst, 1996), particularly the vacancies for university graduates (Figure 2.13).

Age had a positive and significant coefficient at 1 per cent for both periods, indicating that the older aged are likely to be more overeducated relative to matched individuals. This is in contrast to existing models of labour mobility and previous literature, such as Boll *et al.* (2016) and Flisi *et al.* (2014). However, the marginal effect of age was relatively small, even nearly 0 per cent. In other words, the relative probability of overeducation was the same for any age. Meanwhile, the coefficient of age square in the regression was negative and significant, implying that the function is an inverted U-shape. The distribution of age and sample frequency for overeducation is shown in Appendix XIII. Nevertheless, the marginal effect cannot be estimated because the second order derivative is required for interaction/square terms. Moreover, the effect is slightly puzzling. It could be because many people up to the age of early 30s are still looking for appropriate jobs, for example, workers with university qualifications who may have accepted jobs with diploma (lower degree) qualifications (Farooq, 2016). Another possible explanation is because the maximum age limit to apply for jobs in the public sector is 35 years, according to the regulation from Minister of State Apparatus. Some people may prefer to find a job in the private sector first and then apply for a job in the public sector before reaching the maximum age limit of 35 years. Up to certain age, workers possess not only enough education qualifications but also more appropriate experience for their job as their age increases and the probability of overeducation decreases. The downward trend afterwards is consistent with the existing models of labour mobility and with previous literature.

With regards to work-related and firm size variables, the coefficient of tenure was negative and significant for overeducation mean and mode models for 2000. The marginal effect is at -1 per cent. This suggests a negative relationship between tenure and probability of being overeducated or the relative probability of overeducation is, on average, about 1 percentage points lower for one additional year of the job tenure, which agrees with the human capital theory and the job competition theory, as explained in Section 4.2.2. This later turned insignificant in 2014 for both models. Tenure function is a linear function because the coefficient of tenure square was insignificant.

Regarding the part-time variable, the result indicates that coefficients in the main model were negative and significant at 1 per cent for both periods with the marginal effect being -0.07 in 2000 and -0.05 in 2014. The result is in contrast to literatures asserting that part-time work leads to a higher probability of being overeducated (Frank, 1978; Ofek and Merrill, 1997; and Sloane *et al.,* 1999). Cautious interpretation is needed in this case. This finding could be related to one of the (non-financial) incentives of being a public sector worker or to the fact that several private sectors in Indonesia offer flexible working hours (UNDP, 2014; Tjahjono, 2017; and Anell and Hartmann, 2007). Thus, there will be more opportunities for these workers to be with their families or to have an additional job and in turn earning more money for a living.

The next potential variable is private (job sector). The result shows that the coefficient of the private variable was negative and significant at 1 per cent for both models and in both periods. The marginal effects of main model were -0.15 and -0.13 for 2000 and 2014, respectively. This result agrees with Dolton and Vignoles (2000) and Ortiz (2010). McLeod (2006) also adds that all parts of the Indonesian bureaucracy have a very rigid organisational structure in which the number of positions at each level in the hierarchy is fixed mechanically by formula, rather than by reference to the volume of work required to be carried out at that level. Consequently, there is a ubiquitous gross overstaffing at the lower levels. Furthermore, the private sector follows the market mechanism more efficiently in terms of productive and allocative efficiency (Rao, 2015). As such, overeducation is less likely to occur in the private sector.

In terms of the industry variable, manufacturing (wholesale, retail, restaurants and hotels) and social services have a negative and significant coefficient. This suggests that the probability of being overeducated is lower in these industries relative to the agriculture industry. This finding is in line with Allen's (2016) that occupation mismatch (overeducation) in Indonesia tends to be associated with the low education levels of production workers and agriculture labourers. Meanwhile, financial industries have a higher probability for their workers to be overeducated (relative to the agriculture industry) in both periods, which is similar to the finding in Europe that the finance industry is significantly more open to overeducation than manufacturing (Tarvid, 2015).

For firm size variables, the result shows that the coefficient was positive and significant at 1 per cent with respect to small firms (1-19 workers) in both periods. As Dolton and Vignoles (2000) assert, overeducation is the highest amongst those who work for small firms, although generally the incidence of overeducation does not decrease linearly with firm size. Interestingly, more than 70 per cent of the graduates who are overeducated and working in small firms claimed to require no qualifications for their job. It is worth noting that the sample in the present study is concentrated in small firms (around 69.6 per cent and 53 per cent of the total sample in 2000 and 2014, respectively). Thus, the lack of benchmark jobs and formal qualifications which cause a higher incidence of overeducation to be recorded amongst graduates who work in small firms could also occur in the present study.

Regarding urban and capital residency, the result shows a positive and significant coefficient. Hence, living in urban areas increases the likelihood of being overeducated. The marginal effect of urban in 2000 was 0.10, which then decreased slightly to 0.08 in 2014. Statistics Indonesia adds that unemployment rate in urban areas is far higher than in rural areas, for example in February 2017 the unemployment rate in urban areas was 6.5 per cent and in rural areas was 4 per cent (Kontan, 2018). This is in line with Clark *et al.* (2012) who argue that areas with low unemployment rates which have a labour market with relatively lower levels of search frictions make it easier for individuals to sort into an occupation with a required level of education that matches theirs. Besides, more than

60 per cent of the sample in the present study live in urban areas which have higher proportion of overeducation.

Finally, the last variable is capital city. It is interesting that the coefficient changes from positive and significant in 2000 to negative and significant in 2014. The hypothesis of the variable is that workers in urban areas or the capital province (which has high unemployment rate) have a higher risk of being overeducated. One of the possible reasons is improvement in the economy goes in line with rapid technological change. DKI Jakarta is the capital city as well as the centre of the country's politics and businesses. Most headquarter offices are strategically located there, thus the need for workers with higher education qualifications increases recently in line with the rapid development in technological changes in the labour market (Employment Policy Forum, 2018).

4.4.3 Estimation Result by Gender

The present study separates the analysis by gender since gender disparity occurs in Indonesia. The disparity could occur in terms of education levels and wages. In terms of education level, females have a slightly higher average of education attainment than males. The average years of schooling of females in 2000 was 9.73 years, compared to 9.68 years for males. In 2014, the average was 11.83 years for females and 11.33 years for males. The return to education for females is higher than for males, particularly for senior high school and university graduates. This finding is inconsistent with Becker's employer taste model (Becker, 1971). This is possibly because of self-selection: a more limited supply of skilled female workers and different technological requirements in female-dominated and male-dominated jobs (Chapter 3 finding). In terms of the mismatch proportion (Table 4.6), an increase in undereducation and a decrease in overeducation occur when the sample is separated by gender. Yet, both females and males generally experience the same trend.

In terms of mismatch determinants (Table 4.11), this section will highlight several variables related to gender differences: marital status and the presence of young children. It is worth noting that besides those variables, the model includes ethnicity, age, job tenure, job status, industries, and area variables as the control variables in order to limit omitted variable bias.

In terms of undereducation, married and cohabitate status and the other marital status' coefficients are positive and significant, implies workers with those marital statuses have higher probability to be undereducated. This result is in line with the main model. In contrast, for overeducation, married and cohabitate status' coefficients are negative and significant for males and females in both periods. In other words, both married workers have lower probability to be overeducated. In similar vein, McGoldrick and Robst (1996) reject the hypothesis of differential over-qualification; instead, it appears that the large number of vacancies in large labour markets is offset by a larger number of job searchers. With regards to job vacancy and the number of job seekers, Figure 2.13 shows that the number of vacancies for university graduates slightly exceeds the number of job seekers in 2011.

Furthermore, female workers who have young children had lower probability to be undereducated in 2000 and 2014; and had also lower probability to be overeducated in 2000. Schaner and Das (2016) points out that many Indonesian women exit waged work due to family and childcare constraints. As such, a possible explanation for the negative relationship between the presence of young children in the household and the likelihood of undereducation/overeducation is because these workers prefer to exit the waged work rather than being overeducated or undereducated.

4.4.4 Estimation Result by Sector

Similar to gender analysis, sectoral disparity also occurs in Indonesia. One of the key findings of Chapter 3 is that the public sector has a higher return to education than the private sector. The average education level in the public sector is significantly higher than in the private sector in both periods. In 2000, the average education in the public and private sectors were 12.71 years and 9.14 years, respectively. In 2014, the average education in the public sector became 14.64 years while the private sector only reaches 10.92 years. The public sector also has substantially longer job tenure and older workers (on average) than the private sector. Moreover, the public sector has had many changes in regulations related to the education requirements for civil servants (the government regulation 11/2017[54]) and certification of professions (the government regulation

[54] The government regulation of 11/2017 regulates the requirements for civil servants' applicants, such as ageing 18 – 35 years old and having coherent education background with the job requirements.

74/2008). Thus, the determinants could be different. The present study will discuss the difference further in this part.

In terms of the education mismatch proportion, the private sector retains a similar pattern with all individuals. Meanwhile, the public sector has a contrary result; there is an increase in overeducation and a decrease in undereducation. The existing requirements or regulations can be an explanation for this. For instance, the Government Regulation 30/2015 states entry ranks are mainly determined by education level, and increases in rank are largely driven by seniority, with the maximum rank depending on the entry-level of the civil servant. More specifically, an undergraduate entry-level is IIIA and a master's is IIIB. In terms of wages, IIIB with 0-year experience receives a slightly higher wage (around GBP 7 per month) compared to IIIA with 0-year experience. Thus, workers with master's degrees still have an incentive (higher wages) to apply for the same position compared to those with undergraduate degree.

Table 4.12 shows several variables that determine undereducation in the public sector: sex and ethnicity. For overeducation in 2000, there were some determinants such as sex, marital status, age and age square, part time, industry dummy (social services) as well as urban and capital residence. In 2014, personal characteristic variables seem to no longer affect the determinants, but only several industry dummies, firm size and urban residence variables. As the public sector has its own mechanism of remuneration, recruitment and promotion (explained in Chapter 2), thus other factors (outside variables observed in the present study) may have a bigger effect on determining the undereducation and overeducation. Meanwhile, determinants of education mismatch in the private sector are similar throughout all individuals, as the sample is dominated by private sector workers (more than 80 per cent of total sample). Again, the present study finds that the determinants of overeducation are very sensitive in response to sector analysis.

Table 4.11: Determinants of Education Mismatch by Gender, 2000 and 2014

Gender	Males								Females							
	2000				2014				2000				2014			
	Coef.	SE	P>z	ME	Coef.	SE	P>z	ME	Coef.	SE	P>z	ME	Coef.	SE	P>z	ME
UE																
Married and cohabitate	0.284	0.166	*	0.03	0.183	0.120		0.04	0.331	0.218		0.04	0.871	0.160	***	0.16
Other (Separated, divorced and widowed)	0.541	0.441		0.06	0.179	0.247		0.05	0.526	0.293	*	0.07	0.923	0.220	***	0.17
Ethnicity (1=Javanese)	-0.254	0.104	**	-0.02	0.061	0.073		0.01	0.171	0.155		0.02	-0.094	0.092		-0.01
Age	-0.171	0.042	***	0.00	-0.072	0.031	**	-0.00	-0.028	0.058		0.00	-0.068	0.039	*	-0.00
Age squared	0.002	0.001	***		0.001	0.000	***		0.000	0.001			0.001	0.001	**	
Young children (0-5 years old)	0.124	0.131		0.02	-0.114	0.068	*	-0.02	-0.324	0.215		-0.02	-0.176	0.093	*	-0.03
Tenure	0.025	0.022		0.00	-0.006	0.016		-0.00	0.063	0.030	**	0.01	-0.016	0.019		-0.00
Tenure squared	0.000	0.001			-0.000	0.001			-0.001	0.001			0.000	0.001		
Status: part time	-0.109	0.160		-0.00	-0.250	0.114	**	-0.03	0.745	0.181	***	0.08	0.219	0.115	*	0.05
Sector: private	-0.042	0.156		-0.02	0.629	0.141	***	0.13	-0.012	0.247		-0.01	0.929	0.184	***	0.16
Industry2: mining and quarrying	2.942	1.025	***	0.27	1.521	0.241	***	0.26	-8.627	1125.71		-0.85	1.319	0.975		0.20
Industry3: manufacturing	4.257	0.720	***	0.41	1.415	0.163	***	0.26	4.044	1.025	***	0.38	2.244	0.240	***	0.37
Industry4: electricity, gas and water	5.757	0.816	***	0.58	1.009	0.314	***	0.19	-8.777	982.738		-0.84	-11.43	595.72		-1.82
Industry5: construction	1.509	0.821	*	0.14	-0.898	0.270	***	-0.15	4.198	1.267	***	0.38	0.488	0.821		0.04
Industry6: wholesale, retail, restaurants and hotels	5.536	0.720	***	0.57	1.148	0.166	***	0.20	6.319	1.023	***	0.61	1.681	0.240	***	0.27
Industry7: transportation, storage, and communications	2.861	0.783	***	0.26	1.388	0.210	***	0.24	3.862	1.473	***	0.36	1.280	0.630	**	0.20

Industry8: Finance, insurance, real estate and business services	4.431	0.820	***	0.43	0.510	0.211	**	0.08	3.252	1.450	**	0.30	1.009	0.352	***	0.12
Industry9: Social services	4.411	0.717	***	0.45	0.904	0.161	***	0.15	4.022	1.019	***	0.39	0.511	0.232	**	0.07
Firm size2: 20-99 people	-0.566	0.137	***	-0.06	-0.607	0.090	***	-0.10	-0.248	0.190		-0.03	-0.226	0.122	*	-0.05
Firm size3: >= 100 people	-0.873	0.212	***	-0.09	-0.843	0.103	***	-0.15	0.233	0.241		0.01	-0.236	0.134	*	-0.05
Urban	-0.179	0.115		-0.03	-0.254	0.082	***	-0.06	-0.233	0.167		-0.03	-0.467	0.103	***	-0.09
Capital	0.184	0.156		0.01	0.112	0.126		0.03	-0.436	0.225	*	-0.04	-0.130	0.176		-0.02
Constants	-2.677	0.959	***		-1.218	0.521	**		-6.013	1.358	***		-2.445	0.640	***	
OE																
Married and cohabitate	-0.265	0.110	**	-0.06	-0.306	0.123	**	-0.05	-0.417	0.155	***	-0.07	-0.759	0.156	***	-0.12
Other (Separated, divorced and widowed)	-0.340	0.329		-0.08	-0.561	0.305	*	-0.09	-1.267	0.299	***	-0.20	-0.812	0.259	***	-0.13
Ethnicity (1=Javanese)	-0.094	0.076		-0.01	0.062	0.077		0.01	-0.200	0.121	*	-0.03	-0.064	0.104		-0.01
Age	0.118	0.032	***	0.00	0.085	0.035	**	0.00	0.257	0.055	***	0.00	0.207	0.048	***	0.03
Age squared	-0.002	0.000	***		-0.001	0.000	**		-0.004	0.001	***		-0.003	0.001	***	
Young children (0-5 years old)	-0.172	0.099	*	-0.04	0.035	0.067		0.01	-0.530	0.179	***	-0.07	0.004	0.096		0.01
Tenure	-0.061	0.016	***	-0.01	-0.027	0.017	*	-0.00	0.005	0.031		0.00	0.012	0.024		0.00
Tenure squared	0.001	0.001	**		0.000	0.001			-0.002	0.001			-0.001	0.001		
Status: part time	-0.331	0.121	***	-0.06	-0.305	0.119	**	-0.04	-0.458	0.171	***	-0.08	-0.516	0.132	***	-0.07
Sector: private	-0.784	0.121	***	-0.15	-0.786	0.113	***	-0.14	-0.785	0.185	***	-0.11	-0.414	0.136	***	-0.08
Industry2: mining and quarrying	0.572	0.306	*	0.03	-0.292	0.266		-0.10	2.212	1.384		0.48	0.203	0.931		0.00
Industry3: manufacturing	-0.156	0.122		-0.15	-0.938	0.159	***	-0.19	0.204	0.201		0.04	-0.851	0.275	***	-0.17
Industry4: electricity, gas and water	-0.946	0.543	*	-0.34	-0.708	0.358	**	-0.14	1.057	0.950		0.31	1.275	1.033		0.46

Industry5: construction	0.068	0.128		-0.03	0.090	0.169		0.05	1.036	0.435	**	0.08	1.353	0.523	***	0.16
Industry6: wholesale, retail, restaurants and hotels	-1.344	0.192	***	-0.41	-0.509	0.157	***	-0.12	-0.934	0.265	***	-0.25	-0.242	0.259		-0.08
Industry7: transportation, storage, and communications	0.537	0.147	***	0.03	-0.531	0.233	**	-0.13	0.283	0.553		-0.03	0.234	0.557		-0.00
Industry8: Finance, insurance, real estate and business services	-0.069	0.294		-0.13	0.200	0.165		0.11	0.162	0.402		0.04	1.344	0.289	***	0.14
Industry9: Social services	-0.827	0.122	***	-0.28	-0.227	0.137	*	-0.07	-0.724	0.204	***	-0.18	0.266	0.235		0.02
Firm size2: 20-99 people	0.205	0.090	**	0.05	0.152	0.089	*	0.04	0.261	0.137	*	0.04	0.357	0.120	***	0.05
Firm size3: >= 100 people	0.293	0.132	**	0.08	0.447	0.102	***	0.10	0.647	0.173	***	0.09	0.673	0.146	***	0.10
Urban	0.535	0.083	***	0.11	0.464	0.092	**	0.08	0.593	0.147	***	0.09	0.724	0.131	***	0.10
Capital	0.414	0.117	***	0.07	-0.299	0.137	***	-0.05	0.213	0.159		0.04	-0.031	0.180		-0.00
Constants	-1.163	0.492	**		-1.809	0.570			-3.663	0.824	***		-4.623	0.792	***	
Number of obs.	4204				5155				2181				3334			
LR chi2(44)	1042.6				705.1				677				848.08			
Prob > chi2	0				0				0				0			
Pseudo R2	0.1287				0.070				0.1736				0.1306			

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table 4.12: Determinants of Education Mismatch by Sector, 2000 and 2014

	Private								Public							
	2000				2014				2000				2014			
Gender	**Coef.**	**SE**	**P>z**	**ME**	**Coef.**	**SE**	**P>z**	**ME**	**Coef.**	**SE**	**P>z**	**ME**	**Coef.**	**SE**	**P>z**	**ME**
UE																
Sex (1=Female)	-0.447	0.100	***	-0.03	0.112	0.063	*	0.02	-0.574	0.225	**	-0.05	-0.523	0.198	***	-0.05
Married and cohabitate	0.303	0.135	**	0.04	0.413	0.098	***	0.09	0.795	0.548		0.11	0.155	0.414		0.02
Other (Separated, divorced and widowed)	0.565	0.234	**	0.07	0.435	0.161	***	0.10	1.214	0.714	*	0.19	0.447	0.602		0.04
Ethnicity (1=Javanese)	0.072	0.097		0.01	-0.026	0.060		-0.00	-0.878	0.203	***	-0.09	0.092	0.195		0.00
Age	-0.104	0.037	***	0.00	-0.069	0.025	***	0.00	-0.117	0.125		0.00	0.128	0.108		0.00
Age squared	0.001	0.001	**		0.001	0.000	***		0.002	0.002			-0.001	0.001		
Young children (0-5 years old)	-0.039	0.126		0.00	-0.126	0.056	**	-0.02	-0.088	0.241		-0.00	-0.300	0.186		-0.02
Tenure	0.040	0.020	**	0.00	-0.012	0.013		0.00	-0.006	0.043		-0.00	-0.075	0.038	*	-0.00
Tenure squared	0.000	0.001			0.000	0.013			0.000	0.001			0.002	0.001		
Status: part time	0.327	0.130	**	0.04	0.008	0.085		0.01	-0.088	0.262		-0.01	-0.287	0.218		-0.00
Industry2: mining and quarrying	2.893	0.937	***	0.25	1.499	0.236	***	0.27	15.467	7994.62		1.20	1.383	0.774	*	0.14
Industry3: manufacturing	4.130	0.590	***	0.39	1.841	0.138	***	0.35	14.789	693.23		1.68	0.652	0.798		0.11
Industry4: electricity, gas and water	5.619	0.754	***	0.53	1.124	0.314	***	0.21	16.821	693.23		1.92	-0.531	1.190		-0.03
Industry5: construction	1.801	0.683	***	0.16	-0.656	0.252	***	-0.13	0.421	1546.01		0.02	-13.047	683.37		-1.16
Industry6: wholesale, retail, restaurants and hotels	5.799	0.589	***	0.57	1.390	0.140	***	0.26	16.775	693.227		1.90	2.041	1.001	**	0.23
Industry7: transportation, storage, and communications	2.690	0.686	***	0.24	1.533	0.193	***	0.28	15.242	693.227		1.71	-12.107	444.23		-1.12

Industry8: Finance, insurance, real estate and business services	4.130	0.702	***	0.39	0.736	0.184	***	0.11	15.349	693.228		1.71	-0.648	1.163		-0.10
Industry9: Social services	4.209	0.588	***	0.41	0.750	0.136	***	0.13	15.404	693.227		1.77	-0.010	0.508		-0.02
Firm size2: 20-99 people	-0.203	0.122	*	-0.03	-0.471	0.077	***	-0.09	-1.424	0.279	***	-0.16	-0.280	0.204		-0.03
Firm size3: >= 100 people	-0.228	0.163		-0.04	-0.640	0.083	***	-0.13	-1.375	0.495	***	-0.14	-0.000	0.280		-0.03
Urban	-0.197	0.106	**	-0.03	-0.407	0.067	***	-0.09	-0.109	0.207		-0.03	0.041	0.198		0.01
Capital	-0.027	0.133		-0.01	0.023	0.103		0.01	0.101	0.472		0.02	-0.057	0.570		-0.01
Constants	-4.033	0.793	***		-0.965	0.403	**		-14.733	693.230			-3.978	1.94	**	
OE																
Sex (1=Female)	-0.401	0.078	***	-0.06	0.040	0.076		0.00	-0.392	0.173	**	-0.05	-0.178	0.132		-0.02
Married and cohabitate	-0.309	0.094	***	-0.06	-0.571	0.105	***	-0.09	-0.596	0.302	**	-0.13	-0.155	0.250		-0.03
Other (Separated, divorced and widowed)	-0.845	0.226	***	-0.16	-0.790	0.218	***	-0.12	-1.531	0.638	**	-0.31	-0.097	0.432		-0.03
Ethnicity (1=Javanese)	-0.120	0.070	*	-0.02	-0.045	0.069		-0.00	-0.222	0.159		-0.01	0.111	0.137		0.02
Age	0.187	0.030	***	0.00	0.147	0.031	***	0.02	0.167	0.090	*	0.00	0.077	0.074		0.00
Age squared	-0.003	0.000	***		-0.002	0.000	***		-0.002	0.001	*		-0.000	0.000		
Young children (0-5 years old)	-0.248	0.095	***	-0.04	0.064	0.063		0.01	-0.278	0.207		-0.05	-0.083	0.112		-0.01
Tenure	-0.048	0.016	***	-0.01	-0.001	0.017		-0.00	-0.037	0.036		-0.01	-0.048	0.029	*	-0.01
Tenure squared	0.001	0.001			-0.001	0.001			0.000	0.001			0.000	0.001		
Status: part time	-0.293	0.106	***	-0.06	-0.202	0.101	*	-0.03	-0.719	0.251	***	-0.12	-0.883	0.169	***	-0.16
Industry2: mining and quarrying	0.653	0.308	**	0.05	-0.112	0.278		-0.06	16.73	3832.79		2.41	-0.378	0.665		-0.12
Industry3: manufacturing	-0.046	0.108		-0.10	-0.823	0.145	***	-0.17	-0.480	0.431		-0.56	-1.578	0.780	**	-0.32
Industry4: electricity, gas and water	-0.396	0.578		-0.20	-0.506	0.387		-0.10	-0.690	0.779		-0.67	-0.680	0.710		-0.11
Industry5: construction	0.207	0.125	*	0.00	0.417	0.162	***	0.08	0.777	0.552		0.12	-1.717	1.174		-0.08

Industry6: wholesale, retail, restaurants and hotels	-1.229	0.160	***	-0.35	-0.391	0.138	***	-0.10	-0.410	0.622		-0.62	-1.439	1.220		-0.34
Industry7: transportation, storage, and communications	0.662	0.147	***	0.06	-0.393	0.227	*	-0.10	0.191	0.467		0.46	-0.058	0.699		-0.37
Industry8: Finance, insurance, real estate and business services	-0.109	0.261		-0.11	0.490	0.151	***	0.03	0.454	0.622		0.42	0.973	0.477		0.21
Industry9: Social services	-0.774	0.113	***	-0.23	-0.022	0.126		-0.03	-1.130	0.287	***	-0.70	-0.590	0.309	**	-0.11
Firm size2: 20-99 people	0.329	0.084	***	0.06	0.129	0.083	***	0.05	-0.166	0.172		0.02	0.191	0.143		0.05
Firm size3: >= 100 people	0.625	0.113	***	0.12	0.274	0.096	***	0.08	-0.383	0.288		-0.02	0.873	0.179	***	0.17
Urban	0.545	0.079	***	0.10	0.437	0.088	***	0.08	0.616	0.185	***	0.11	0.559	0.145	***	0.11
Capital	0.260	0.099	***	0.05	-0.256	0.116	**	-0.03	1.054	0.296	***	0.18	0.214	0.326		0.04
Constants	-2.944	0.441	***		-3.710	0.499	***		-2.207	1.507			-1.502	1.296		
Number of obs.	5379				7084				1006				1405			
LR chi2(44)	1497.9				873.14				276.24				217.73			
Prob > chi2	0				0				0				0			
Pseudo R2	0.1482				0.071				0.1405				0.0840			

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

4.5 The Sensitivity Test of the Results

4.5.1 Multinomial Logit/MNL Model (Mean)

The first sensitivity test is conducted by comparing the mode and mean results, as most studies use mean as a base of their analysis (see Sicherman, 1991; Kiker *et al.,* 1997; Chevalier, 2000; Nazara and Safuan, 2005; and ILO, 2017b). Mean and mode are different by definition, as mean refers to the average value or the sum of all of the given data which is then divided by the number of data entry or observation while mode refers to the number that occurs most often in the category. Commonly the mean, median, and mode of a normal distribution are equal. However, the distributions of educational attainment are potential non-normally distributed, this could indicate there is a case for looking at the different measures of central tendency in the defining over-education.

In terms of the mismatch determinants, it seems that the set of variables are sensitive to the method used, as different method gives different result of the determinants (Table 4.13). However, most of the variables are still significant in determining education mismatch in Indonesia; some variables which have different results are sex, ethnicity, age and age square, young children, and part time (job status).

Table 4.13: Determinants of Education Mismatch, MNL (Mean), 2000 and 2014

	2000				2014			
Mean	**Coef.**	**SE**	**P>z**	**ME**	**Coef.**	**SE**	**P>z**	**ME**
UE								
Sex (1=Female)	-0.237	0.092	***	-0.02	0.295	0.070	***	0.04
Married and cohabitate	0.216	0.130	*	0.03	0.388	0.116	***	0.06
Other (Separated, divorced and widowed)	0.468	0.214	**	0.07	0.657	0.171	***	0.09
Ethnicity (1=Javanese)	-0.056	0.084		-0.00	-0.193	0.066	***	-0.02
Age	0.002	0.034		0.00	-0.023	0.028		-0.01
Age squared	0.000	0.000			0.001	0.000	***	
Young children (0-5 years old)	0.015	0.106		0.01	-0.160	0.063	**	-0.02
Tenure	0.039	0.017	**	0.00	-0.026	0.013	*	0.00
Tenure squared	0.000	0.001			0.001	0.000		
Status: part time	0.112	0.113		0.02	0.066	0.087		0.01
Sector: private	0.216	0.127	*	0.04	1.715	0.154	***	0.22
Industry2: mining and quarrying	3.343	0.661	***	0.07	0.873	0.282	***	0.11
Industry3: manufacturing	3.643	0.461	***	0.12	1.532	0.151	***	0.20
Industry4: electricity, gas and water	4.752	0.605	***	0.27	0.788	0.367	**	0.11
Industry5: construction	1.245	0.575	**	0.01	2.228	0.187	***	0.24
Industry6: wholesale, retail, restaurants and hotels	4.739	0.461	***	0.30	0.942	0.155	***	0.12
Industry7: transportation, storage, and communications	5.218	0.470	***	0.32	1.020	0.225	***	0.13

Industry8: Finance, insurance, real estate and business services	3.428	0.590	***	0.10	0.301	0.220		0.03
Industry9: Social services	3.937	0.457	***	0.16	0.794	0.148	***	0.10
Firm size2: 20-99 people	-0.340	0.107	***	-0.04	-0.481	0.084	***	-0.06
Firm size3: >= 100 people	-0.611	0.159	***	-0.06	-0.937	0.101	***	-0.12
Urban	-0.322	0.090	***	-0.05	-0.460	0.072	***	-0.06
Capital	-0.094	0.127		-0.01	-0.172	0.124		-0.01
Constants	-5.833	0.694	***		-4.220	0.472	***	
OE								
Sex (1=Female)	-0.276	0.083	***	-0.03	-0.111	0.070		-0.02
Married and cohabitate	-0.428	0.103	***	-0.06	-0.596	0.101	***	-0.08
Other (Separated, divorced and widowed)	-1.064	0.266	***	-0.13	-0.691	0.206	***	-0.10
Ethnicity (1=Javanese)	-0.202	0.074	***	-0.03	-0.019	0.065		0.00
Age	0.318	0.033	***	0.00	0.128	0.029	***	0.02
Age squared	-0.005	0.000	***		-0.002	0.000	***	
Young children (0-5 years old)	-0.277	0.103	***	-0.04	0.032	0.058		0.01
Tenure	-0.056	0.016	***	-0.01	-0.006	0.014		0.00
Tenure squared	0.001	0.001			-0.001	0.001		
Status: part time	-0.583	0.126	***	-0.07	-0.334	0.095	***	-0.04
Sector: private	-0.866	0.110	***	-0.13	-0.809	0.092	***	-0.13
Industry2: mining and quarrying	0.930	0.339	***	0.14	-0.194	0.249		-0.04
Industry3: manufacturing	0.060	0.130		-0.01	-0.862	0.142	***	-0.13
Industry4: electricity, gas and water	0.154	0.447		-0.03	-0.649	0.362	*	-0.09
Industry5: construction	0.307	0.148	**	0.05	1.188	0.173	***	0.10
Industry6: wholesale, retail, restaurants and hotels	-0.835	0.177	***	-0.12	-0.472	0.139	***	-0.08
Industry7: transportation, storage, and communications	0.906	0.176	***	0.06	-0.508	0.222	**	-0.08
Industry8: Finance, insurance, real estate and business services	-0.024	0.265		-0.02	0.193	0.151		0.02
Industry9: Social services	-0.292	0.127	**	-0.06	-0.265	0.124	**	-0.05
Firm size2: 20-99 people	0.272	0.086	***	0.04	0.328	0.076	***	0.05
Firm size3: >= 100 people	0.407	0.117	***	0.07	0.582	0.086	***	0.09
Urban	0.790	0.089	***	0.10	0.471	0.079	***	0.07
Capital	0.194	0.103	*	0.03	-0.385	0.120	***	-0.04
Constants	-5.530	0.527	***		-2.929	0.485	***	
Number of obs.	6385				8489			
LR chi2(46)	1270.34				1465.26			
Prob > chi2	0				0			
Pseudo R2	0.1222				0.1017			

Source: The author’s calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

4.5.2 Multinomial Probit/MNP Model

To test the sensitivity of the results, the present study carries out the multinomial probit (MNP) model and compares the results with those from MNL. Table 4.14 shows the results of MNP for both periods. By comparing Table 4.14 and Table 4.10 it is apparent that the patterns of coefficient variables for both undereducation and overeducation are similar in the significance, except for several variables such as urban (UE) and ethnicity (UE). In terms of the marginal effects, the values are relatively similar for both MNL and MNP models. Hence, the interpretations of the results are similar for MNL and MNP. The results also show that MNL estimation performs as well as (if not better than) MNP; and is in line with the previous study such as Dow and Endersby (2004). Therefore, the conclusion in general is the same: the mismatch is determined by workers' and job's characteristics (the Assignment Model).

Table 4.14: Determinants of Education Mismatch: MNP (Mode), 2000 and 2014

	2000				2014			
MNP	**Coef.**	**SE**	**P>z**	**ME**	**Coef.**	**SE**	**P>z**	**ME**
UE								
Sex (1=Female)	-0.377	0.067	***	-0.04	0.017	0.047		0.01
Married and cohabitate	0.195	0.095	**	0.03	0.295	0.073	***	0.08
Other (Separated, divorced and widowed)	0.404	0.167	**	0.08	0.333	0.125	***	0.09
Ethnicity (1=Javanese)	-0.109	0.063	*	-0.01	0.004	0.045		0.00
Age	-0.077	0.025	***	-0.00	-0.039	0.019	**	-0.00
Age squared	0.001	0.000	***		0.001	0.000	***	
Young children (0-5 years old)	-0.011	0.081		-0.01	-0.114	0.042	***	-0.03
Tenure	0.022	0.013	*	0.00	-0.014	0.009		-0.00
Tenure squared	0.000	0.000			0.000	0.000		
Status: part time	0.126	0.087		0.03	-0.086	0.062		-0.00
Sector: private	-0.087	0.096		-0.02	0.477	0.079	***	0.12
Industry2: mining and quarrying	1.566	0.501	***	0.03	1.116	0.178	***	0.23
Industry3: manufacturing	2.326	0.257	***	0.12	1.279	0.100	***	0.30
Industry4: electricity, gas and water	3.311	0.396	***	0.37	0.696	0.233	***	0.14
Industry5: construction	0.939	0.305	***	0.01	-0.484	0.164	***	-0.05
Industry6: wholesale, retail, restaurants and hotels	3.517	0.258	***	0.46	0.947	0.100	***	0.19
Industry7: transportation,	1.604	0.305	***	0.03	1.014	0.148	***	0.21

storage, and communications								
Industry8: Finance, insurance, real estate and business services	2.282	0.355	***	0.11	0.486	0.130	***	0.05
Industry9: Social services	2.312	0.254	***	0.16	0.474	0.096	***	0.08
Firm size2: 20-99 people	-0.260	0.078	***	-0.04	-0.340	0.056	***	-0.08
Firm size3: >= 100 people	-0.240	0.109	**	-0.05	-0.415	0.063	***	-0.11
Urban	-0.095	0.070		-0.03	-0.232	0.051	***	-0.07
Capital	-0.001	0.092		-0.01	0.001	0.080		0.01
Constants	-2.025	0.455	***		-1.376	0.312	***	
OE								
Sex (1=Female)	-0.341	0.056	***	-0.06	-0.028	0.050		-0.01
Married and cohabitate	-0.232	0.074	***	-0.06	-0.345	0.074	***	-0.09
Other (Separated, divorced and widowed)	-0.656	0.162	***	-0.15	-0.464	0.142	***	-0.11
Ethnicity (1=Javanese)	-0.107	0.052	**	-0.02	0.009	0.047		0.01
Age	0.106	0.021	***	0.00	0.090	0.021	***	0.02
Age squared	-0.002	0.000	***		-0.001	0.000	***	
Young children (0-5 years old)	-0.216	0.068	***	-0.05	0.018	0.042		0.01
Tenure	-0.035	0.011	***	-0.01	-0.009	0.010		-0.00
Tenure squared	0.000	0.000			0.000	0.000		
Status: part time	-0.286	0.076	***	-0.07	-0.297	0.066	***	-0.05
Sector: private	-0.646	0.081	***	-0.15	-0.473	0.069	***	-0.13
Industry2: mining and quarrying	0.618	0.260	**	0.15	-0.062	0.192		-0.08
Industry3: manufacturing	0.037	0.085		0.04	-0.576	0.101	***	-0.17
Industry4: electricity, gas and water	-0.316	0.347		-0.19	-0.357	0.252		-0.11
Industry5: construction	0.188	0.102	*	0.04	0.191	0.126		0.06
Industry6: wholesale, retail, restaurants and hotels	-0.743	0.111	***	-0.28	-0.295	0.100	***	-0.11
Industry7: transportation, storage, and communications	0.552	0.119	***	0.13	-0.268	0.160	*	-0.11
Industry8: Finance, insurance, real estate and business services	0.079	0.198		0.02	-0.423	0.113	***	-0.08
Industry9: Social services	-0.595	0.084	***	-0.18	-0.072	0.090		-0.04
Firm size2: 20-99 people	0.191	0.061	***	0.05	0.148	0.055	***	0.05

Firm size3: >= 100 people	0.364	0.086	***	0.10	0.347	0.064	***	0.10
Urban	0.429	0.058	***	0.10	0.383	0.056	***	0.08
Capital	0.276	0.076	***	0.06	-0.154	0.083	*	-0.03
Constants	-1.184	0.331	***		-2.027	0.343	***	
Number of obs.	6385				8489			
Wald chi2(46)	1206.94				1160.54			
Prob > chi2	0				0			

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

4.5.3 Adding Casual Workers

Besides comparing different methods, the sensitivity could also be tested by adding another sample category, *i.e.* casual workers. However, due to the lack of data availability in the 2000 survey, the present study only performs this addition for the data from the 2014 survey. The new model then adds another control variable (casual workers). The hypothesis of casual worker in the model is that the waged sector has better matches than casual worker, as per Chua and Chun's argument (2016).

Table 4.15 shows the sample distribution with casual workers where the number of observations increases from 8,489 to 10,594 individuals. On one hand, adding casual workers into the sample decreases undereducation proportion from 22.75 per cent to 21.8 per cent. On the other hand, the addition increases overeducation proportion from 23.35 per cent to 25.14 per cent. From 2,105 individuals of casual worker, 18 per cent is undereducated and 32.35 per cent is overeducated. Overall, this addition slightly increases the proportion of mismatch (overeducation and undereducation) from 46.1 per cent in 2014 (the main model) to 46.94 per cent in the new model with casual workers added. Thus, adding casual workers data in the model increases the mismatch proportion. The summary statistics of related variables is provided in Appendix XV.

In terms of mismatch determinants (Table 4.16), most of the variables have similar sign of coefficient and significance as well as similar value of marginal effect, except for some variables such as sex and young children presence for undereducation; and several industry dummies for overeducation. The results show that coefficient of casual worker is positive and significant for undereducation as well as negative and significant for overeducation, suggesting that casual workers are more likely to be undereducated, relative to being matched. In this respect, Allen (2016) asserts that the majority of jobs created use short-term contracting arrangements or are in the informal sector. Employment quality is also a major issue and compliance with existing labour regulations is very low. Youth unemployment, skills shortages, and education/skills mismatches are also persistent challenges.

Table 4.15: Sample Distribution with Casual Workers, 2014

Match (Mode)	Freq.	Per cent	Cum.
UE	2,309	21.8	21.8
M	5,622	53.07	74.86
OE	2,663	25.14	100
Total	10,594	100	

Source: The author's calculation.

Table 4.16: Determinants of Education Mismatch, MNL (Mean), 2000 and 2014

	2014			
Casual	**Coef.**	**SE**	**P>z**	**ME**
UE				
Sex (1=Female)	-0.095	0.055	*	-0.01
Married and cohabitate	0.363	0.089	***	0.07
Other (Separated, divorced and widowed)	0.346	0.141	**	0.07
Ethnicity (1=Javanese)	0.006	0.052		0.00
Age	-0.051	0.022	**	-0.00
Age squared	0.001	0.000	***	
Young children (0-5 years old)	-0.070	0.050		-0.01
Tenure	-0.011	0.010		-0.00
Tenure squared	0.000	0.000		
Status: part time	-0.092	0.069		-0.00
Sector: private	0.673	0.107	***	0.12
Industry2: mining and quarrying	0.772	0.226	***	0.11
Industry3: manufacturing	1.611	0.114	***	0.29
Industry4: electricity, gas and water	0.883	0.291	***	0.14
Industry5: construction	-0.805	0.190	***	-0.06
Industry6: wholesale, retail, restaurants and hotels	1.253	0.114	***	0.22
Industry7: transportation, storage, and communications	1.381	0.163	***	0.25
Industry8: Finance, insurance, real estate and business services	0.388	0.166	**	0.04
Industry9: Social services	0.739	0.108	***	0.11
Firm size2: 20-99 people	-0.393	0.069	***	-0.06
Firm size3: >= 100 people	-0.483	0.078	***	-0.08
Urban	-0.121	0.059	**	-0.03
Capital	0.060	0.093		0.02
Casual	0.741	0.122	***	0.13

Constants	-1.748	0.362	***	
OE				
Sex (1=Female)	-0.221	0.056	***	-0.03
Married and cohabitate	-0.296	0.080	***	-0.07
Other (Separated, divorced and widowed)	-0.451	0.148	***	-0.09
Ethnicity (1=Javanese)	0.071	0.050		0.01
Age	0.084	0.021	***	-0.00
Age squared	-0.001	0.000	***	
Young children (0-5 years old)	0.034	0.045		0.01
Tenure	-0.003	0.010		-0.00
Tenure squared	-0.001	0.000	*	
Status: part time	-0.293	0.068	***	-0.04
Sector: private	-0.581	0.081	***	-0.13
Industry2: mining and quarrying	-0.408	0.198	**	-0.11
Industry3: manufacturing	-0.855	0.102	***	-0.22
Industry4: electricity, gas and water	-0.848	0.300	***	-0.19
Industry5: construction	0.578	0.097	***	0.15
Industry6: wholesale, retail, restaurants and hotels	-0.877	0.103	***	-0.21
Industry7: transportation, storage, and communications	-0.994	0.190	***	-0.23
Industry8: Finance, insurance, real estate and business services	0.035	0.115		-0.01
Industry9: Social services	-0.557	0.084	***	-0.14
Firm size2: 20-99 people	0.068	0.062		0.03
Firm size3: >= 100 people	0.314	0.075	***	0.08
Urban	0.240	0.056	***	0.05
Capital	-0.256	0.097	***	-0.04
Casual	-0.355	0.098	***	-0.09
Constants	-1.001	0.353	***	
Number of obs.	10594			
LR chi2(48)	1629.13			
Prob > chi2	0			
Pseudo R2	0.0757			

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

4.6 Conclusion

Throughout the literature on the economics of education, education mismatch is not a new issue in Indonesia but most of previous studies focus only on overeducation. This chapter has focused on answering the research questions *i.e.* does education mismatch (both undereducation and overeducation) exist in the waged sector in Indonesia? What are the estimated proportions of education mismatch in 2000 and 2014? And, how does education mismatch change between these periods?

The results show that mismatch (both overeducation and undereducation) indeed increases in Indonesia in both years, i.e., 2000 and 2014; mismatch is a manifestation of undereducation. Around 53-58 per cent of the sample was in the match category for both periods. In 2000, the proportion of the undereducated was 13.6 per cent, which then increased to 22.8 per cent in 2014. The overeducation proportions in 2000 and 2014 were around 28.5 per cent and 23.4 per cent, respectively. Similar to overeducation, undereducation is inefficient for the economy because an increase in the incidence of undereducation could decrease productivity on average (Kampelmann and Rycx, 2012). Some possible explanations for the occurrence of overeducation/undereducation in Indonesia are firstly, the technological change (Oliveira, 2000) as the country's economy experiences a rapid growth, a modernisation of the industrial structure, and an expansion in higher education. ILO (2017b) argues that the current wave of technological advances in Indonesia is happening at a much faster rate, which undoubtedly will have a major impact on the production processes and the organisation of work. Secondly, it is due to under-production and a high demand for workers with higher education qualifications, as Carnevale and Rose (2013) explains. Unfortunately, the present research is limited by the lack of available labour demand analysis. Figure 2.13 (data of registered job seekers and the number of vacancies) could give an indication of the supply and demand of labour where the number of vacancies for university graduates is higher than the number of job seekers, possibly causing the increase in the demand for workers with higher education qualifications to occur in Indonesia. Thirdly, the minimum wage policy possibly disheartens individuals from pursuing higher educations because the increase of minimum wage has a mixed effect to different groups of workers (Smeru, 2001). And finally, the weakness of the Realise Method (RM), particularly the change in the mode only reflects the change in average workers' education and the measurement cannot observe this change because the jobs demand higher educations.

While Indonesia's labour market experiences rapid technological development, that technology has resulted in a move up from unskilled to semi-skilled employment. In addition, both of vertical and horizontal mismatches occur in Indonesia, as Alisjahbana, *et. al.* (2017) found that most of the workers with science and engineering degrees are considered to be over-educated regardless of major, or only about 25 per cent to 44 per cent of the workers depending on major background and jobs held are considered to be adequately educated.

Furthermore, the highest unemployment rate occurs to senior high school and vocational high school qualifications, and where labour with medium education qualifications is not optimally absorbed by the jobs available. On the demand side, most companies would hardly hire semi-skilled and high-skilled workers (Employment Policy Forum, 2018). Thus, the mismatch occurs not only as education-job mismatch, but also as education-skill-job mismatch. It appears that mitigating these issues require government policies on scholarships for higher education or other education policies aimed at increasing participation in higher educations, improving compulsory education program, developing curriculum that is in line with the industries' need, and promoting public-private collaboration to provide comprehensive and systematic training (internship) programs.

In terms of the estimation technique, detailed education variable and occupation classification (as long as the minimum number of observations is retained: 30 observation for each category) are required to determine undereducation and overeducation objectively. The aims are to minimise heterogeneity and low observation problem in each occupation category. Moreover, estimating mismatch using the mode method is proven better than using the mean method, because mode is used commonly for categorical variable (like mismatch variable). Mode also has better estimates when the data are skewed or not normally distributed.

The other research questions are: what are the variables that determine undereducation and/or overeducation? How does the aggregate trend of education mismatch change between 2000 and 2014? And are there any distinctions among gender and sectors?

In terms of the sample's sub-categories, sector substantially affects the mismatch; the private sector is the main contributor in the increase of undereducation whereas gender seems to have insignificant effect on mismatch. Different trends of mismatch are found in the public sector; there was an increase in overeducation and a decrease in undereducation between 2000 and 2014 periods. It is possible that this is due to the change (adjustment) of regulations in public sector, for instance, the Government Regulation 30/2015, entry ranks are mainly determined by education level, and increases in rank are largely driven by seniority, with the maximum rank depending on the entry-level of the civil servant. Furthermore, the public sector recently prefers to hire workers with high education levels; those with at least senior high school qualifications. For unskilled jobs such as cleaning, however, the public sector prefers outsourcing or using private firms' services rather than hiring directly.

In terms of determinants, the variables are slightly sensitive to the different methods used, the set of variables, the sector/gender and the periods. The results show that education mismatch in Indonesia is determined by personal and household characteristics, work related and firm size as well as area of residency variables which are all observed in this research and which is in line with the Assignment Models (mismatch is determined by workers' and job's characteristics).

Therefore, the research implications are the estimation techniques and occupational categories could affect the analysis. As the assignment theory suggest, mismatch is determined by workers' and job's characteristics. Workers can minimise the incidence of mismatch by considering their characteristics. For instance, working in urban areas where education quality is relatively higher than in rural areas; in urban areas, most of the population have higher education attainment and the services sector demands highly qualified workers. Thus, those working in urban areas are less likely to be undereducated, though they are more likely to be overeducated.

The present study is aware of the limitations in this chapter, such as: the mismatch is not only determined by individual factors and considering the demand for labour could improve the analysis from the labour supply and demand point of view. Information on the education requirements from firms could also provide a better measurement of the changes where the jobs demanded higher education qualifications. Moreover, it is necessary to conduct further study using the data on skills and subjects of study, hence addressing the vertical and horizontal mismatches. In addition, an impact analysis is usually related to issues in wage; this will be discussed further in the following chapter (Chapter 5).

Chapter 5 The effect of Education Mismatch on Wages in the Waged Sector in Indonesia

5.1 Introduction

In spite of the education expansion in Indonesia, education mismatch[55] increases as driven by the increase in under-education (see Chapter 4). There are some possible reasons for this issue. On the supply side, the population with low education attainment is still far higher than those with higher education, meanwhile the quality of the education is still relatively low. On the demand side, the production technology in Indonesia changes rapidly and could affect more companies to prefer workers with higher educational qualifications (Allen, 2016). Allen and Kyloh (2016) argues that young people in Indonesia often leave school and enter the workforce with qualifications that do not match the needs of potential employers due to low education attainment or quality and poor career guidance. With an undereducated workforce, there is a risk of weaker productivity growth and a slower structural transition to higher value-added activities.

Duncan and Hoffman (1981) develop the Over-Required-Undereducated (ORU) model, which is an extension of Mincer wage equation. In ORU model, education variables are decomposed into required years of schooling (REQ), years of deficit schooling or undereducated (UE), and years of surplus schooling or overeducated (OE). Previous empirical studies (such as: Daly et al. (2000) and Korpi and Tahlin (2007)) using Duncan and Hoffman's model find one extra year of surplus schooling (overeducated) gains premium wages, since the coefficient of wage return to one-year surplus schooling is positive and significant. Meanwhile, one extra year of deficit schooling (undereducated) receives penalty wages.

In Indonesia, there are several empirical researchers exploring this issue, yet most of them do not account for unobserved heterogeneity by using panel data analysis. Also, most of them only focus on the analysis on higher education, while one of the main challenges in Indonesia is the domination of workers with lower education levels in the labour market (see Chapter 2). In addition, the model used is the modified Verdugo and Verdugo model (1989). For instance, using just the data from one survey period (Sakernas of 2014),

[55] Education mismatch is measured by objective measures, *i.e.* calculating the mode and standard deviation, or known as realised method (RM). This measurement is in the definition in Chapter 4.

Alisjahbana *et al*. (2017) find that overeducated and undereducated workers receive lower wages compared to those with adequate educational levels for their jobs. This result is inconsistent with Duncan and Hoffman's findings, since the model used is different (the Verdugo and Verdugo model).

The present study contributes to the existing analysis on wages from the aspects of both the undereducation and overeducation in Indonesia, especially considering that undereducation is a rarely studied topic in Indonesia. Similarly, the present study contributes to the international literature on empirical evidence of Duncan and Hoffman's findings in developing countries. Furthermore, the increasing trend of undereducation may decrease productivity on average. Thus, the analysis of education mismatch is a matter of public policy interest. Also, most of the previous studies in Indonesia use cross-section models; as a result, the panel data used in the present study can contribute to enrich the method used to analyse wages and education mismatch in Indonesia.

This chapter explores the relationship between mismatch incidence and wage by focusing on the vertical mismatch among workers in the waged sector by using panel analysis. The aims of this chapter are: to investigate the extent to which mismatch incidences affect wages; considering the unobserved heterogeneity, to see if mismatch incidences still affect wages; to investigate the effect of gender and sector on returns associated with education mismatch; and to contribute to the existing literature on education mismatch and returns by taking into account the influence of unobserved heterogeneity. The research questions are: (5.1) does education mismatch (undereducation and/or overeducation) contribute to determining wages in Indonesia? (5.2) considering unobserved heterogeneity, does education mismatch still contribute to determining wages, and (5.3) do returns associated with education mismatch differ by gender and by sector?

To estimate the wages, this research employs an extension of the Mincer wage equation as proposed by Duncan and Hoffman (1981) *i.e.* the ORU model with pooled OLS to answer research question 5.1, and with panel analysis to answer research question 5.2. There are several advantages of using the Duncan and Hoffman model. Firstly, the model can provide the estimates necessary to evaluate wages from the monetary perspective. Secondly, the model also allows the analysis of both the return to under-education and over-education. Thirdly, the model also has a better interpretation of mismatch-wage relationships, as it allows the analysis of premium wages to each additional surplus of education and penalty wages to each additional deficit of education, not just overeducation and undereducation status. This is important since workers can have various years of overeducation/undereducation, for instance, workers with 1 year surplus of education may have different levels of wage compared to workers with 3 years surplus of education. Fourthly, the panel data model provides a major advantage as it can be used to deal with unobserved heterogeneity problem. Panel data model also accommodates individual effects such as cognitive ability and motivation which are otherwise given and assumed constant over time (Nielsen, 2014).

This chapter uses the data from IFLS3 of 2000, IFLS4 of 2007, and IFLS5 of 2014 to answer the research questions. Adding the 2007 data has the benefit of increasing the number of individuals, so the data become more informative and the estimates become more reliable (Baltagi, 2015). Moreover, adding the 2007 wave narrows down the gap between periods of analysis, thus allowing the present study to analyse around 7 years period of change in wages in Indonesia. The waves were chosen to represent the condition before the education reform period (the data from 2000), during the initial period of the reform (the data from 2007) and after the reform (the data from 2014). To estimate the education mismatch, this research employs objective measures (realised method/RM) by calculating the mode and standard deviation (see Chapter 4).

The present study also extends the analysis based on gender and sector, as previous chapters find that gender and sector have a substantial effect on wages and mismatch incidences. In terms of gender, the return to education for females is higher than that for males (Chapter 3 finding). Yet, gender seems to have an insignificant effect in determining education mismatch (Chapter 4 finding). Thus, it is interesting to further investigate the relationship between gender, mismatch and return associated with education mismatch. For instance, Duncan and Hoffman (1981) find that overeducated females tend to have lower return than males.

In terms of sector, the public sector is very attractive in Indonesia. The ratio of job opportunities in the public sector to the number of applicants is around 1:200 (Sindo, 2013). This is because the public sector pays higher wages than the private sector (Chapter 3 finding). Moreover, Chapter 4 finds that there are different education mismatch trends. Generally, there is a decrease in overeducation and an increase in undereducation in the private sector. In contrast, the public sector experiences an increase in overeducation and a decrease in under-education. Thus, sectors may contribute to the difference in wages between overeducation and undereducation in Indonesia. Dolton and Vignoles (2000) further that overeducated workers in the public sector earn less than the overeducated ones in the private sector due to the relatively less competitive nature of the public sector. The structure of Chapter 5 is organised as follows: Section 5.2 reviews various literature on mismatch and wage, including the persistence of mismatch. Section 5.3 explains the data and methods of the study. Section 5.4 discusses the estimation results. Finally, Section 5.5 presents the conclusion of this chapter.

5.2 Literature Review

Principally, this part will discuss the literature, both the theories and empirical evidence. The discussion is divided into several parts: related theories, the ORU model, empirical studies on mismatch and wage, the dynamic of wage and mismatch, as well as wage and education mismatch based on gender and sector.

5.2.1 Related Theories

Overeducation and undereducation approach provides a much more comprehensive picture of the returns from years of education in the labour market. These have the appeal which links the demand-side considerations into the typical supply-oriented human capital approach to wage determination (Dockery and Miller, 2012). The consequences of mismatch have mostly been addressed in terms of wages in the literature on the topic.

In this part, the present study highlights several theories related to education mismatch and wage. Firstly, the human capital theory suggests that overeducated graduates will earn less than their education as their peers in graduate jobs, *ceteris paribus* (Dolton and Vignoles, 2000). This because overeducation may be a temporary consequence of graduate accepting lower-paid occupations which require fewer skills to increase their experience in the field to improve their chances of obtaining a more suitable job in the long term. Thus, the overeducated should not have any pay penalty inflicted upon them. Variations in the human capital theory relax the assumptions of perfect abilities of firms in adjusting their production technologies to changes in the relative supply of labour instantaneously. This hampers the firms' abilities to fully utilise their workers' education, restricts their productivity and lowers their wages (Cedefop, 2009).

The job competition model of Lester Thurow (Thurow, 1975) assumes a fixed productivity and education (skills) requirement. The choice of production techniques is also assumed not responsive to changes in the relative supply of different education (skills) groups. As a result, when individuals work below their education (skills) level, the individuals will have levels of wage and productivity that are identical to those of individuals whose level of education matches their education (skills) requirement. In other words, Thurow's job competition model suggests that the return to surplus education is zero (Duncan and Hoffman, 1981). Similarly, institutional theory suggests that only job characteristics determine wages (Cedefop, 2009).

The assignment theory and heterogeneous skill theory state that a worker's wage is determined by both the required levels of education and the actual skills possessed. A close link between educational and skill mismatch is assumed in the assignment theory. The theory suggests that overeducated workers are unable to use all of their skills. Hence, they are less productive than similar individuals with jobs for which their educational attainment is appropriate (matched workers). In contrast, heterogeneous skill theory suggests a much weaker link between educational and skill mismatch. The main assumption is that there is a significant variability in terms of skills and (sometimes unobservable) abilities among individuals with the same level of schooling. In fact, the overeducated are less able individuals who actually match with their jobs in terms of skills and abilities. These lower returns are consistent with a scenario in which overeducated workers' jobs are imposing an upper limit on the extent to which they can utilise their skills with this productivity ceiling reflected in lower wages (McGuinness, 2006).

5.2.2 Wage Equation: Required-Under-Overeducated (ORU)

Duncan and Hoffman Model

One of the pioneering studies in measuring wage is conducted by Duncan and Hoffman (1981) by extending the Mincer wage equation[56]. The model is commonly called the standard Over-Required-Undereducated (ORU) specification. In the ORU model, education variables are decomposed into required years of schooling (REQ), years of deficit schooling (UE), and years of surplus schooling (OE). Most empirical research uses

[56] Similar to the model in Chapter 3.

this specification, such as Dolton and Vignoles (2000) and Nieto and Ramos (2017). Hartog (2000) points that Duncan and Hoffman's model lacks a coherent theoretical framework. Nevertheless, the model has some advantages (as explained in Section 5.1) and it has proven itself as an extension of Mincer wage equation by passing statistical testing in several countries for several periods.

Duncan and Hoffman's model for pooled model across time is as follows:

$$\ln W_{i,t} = \beta_0 + \beta_1 X_{i,n,t}^{\mathrm{UE}} + \beta_2 X_{i,n,t}^{\mathrm{OE}} + \beta_3 X_{i,n,t}^{\mathrm{REQ}} + \sum_{k=1}^{Q} \beta_{3+n} \mathrm{K}_{i,n,t} + \varepsilon_{i,t} \quad (5.1)$$

where:

$\ln W_{i,t}$: log of real hourly wage (similar dependent variable as Chapter 3),

$X_{i,n,t}^{REQ}$: years of required schooling (equal to mean or mode if using objective measures) ,

$X_{i,n,t}^{UE}$: years of deficit schooling relative to the average level of education in the individual's occupational category (refers to undereducation),

$X_{i,n,t}^{OE}$: years of surplus schooling relative to the average level of education in the individual's occupational category (refers to overeducation),

K: set or vector of other explanatory (control) variables,

i is individual (i=1…I), n is the number of explanatory variables

β_1: the estimated wage returns to one-year deficit schooling, β_2: the estimated wage returns to one year of surplus schooling, and β_3: the estimated wage returns to one year of required schooling.

By applying the model and using the US data (the Panel Study of Income Dynamics in 1976) with experience and experience-squared as well as city-size as control variables and residence in the south (location) as a dummy variable, Duncan and Hoffman (1981) find that the surplus of schooling coefficient (β_2) is positive and significant and the deficit of schooling coefficient (β_1) is negative and significant on wage. This implies that one extra year of surplus schooling (overeducated) still gains premium wages; while one extra year of deficit schooling (undereducated) receives penalty wages. In terms of coefficient values, Duncan and Hoffman find that wage return to one-year surplus schooling is approximately half as large as the wage return to one-year required schooling. Meanwhile, wage return to one-year deficit schooling is less than wage return to required schooling in absolute value. It is worth noting, the limitation in the interpretation of results obtained from this approach is that a positive coefficient on overeducation could in fact indicate systematic mismeasurement of required years of schooling, this because the issues concerning endogeneity and measurement error, and the estimated return to required/under/overeducation cannot be interpreted as casual (Lauven and Oosterbeek, 2011).

Duncan and Hoffman also add that a plausible explanation for this is the productivity levels on the job are more variable in the US. This finding is also in line with the Human Capital theory's prediction and it also appears to support an assignment model interpretation (Sattinger, 1993)[57]. Meanwhile, Dolton and Vignoles (2000) suggest that the positive return to surplus education may explain why individuals continue to invest in education although a significant number ends up being overeducated. In contrast, this finding rejects Thurow's job competition model (Thurow, 1975) which suggests that only job characteristics, *i.e.* required education levels, determine earnings.

Key Findings of ORU approach

Hartog (2000) provides a review of empirical findings from the ORU approach and a discussion of the methodological issues. Hartog identifies four key findings from this literature: (1) the return from required years of schooling is higher than the return from actual years of schooling that is obtained from the standard Mincer wage equation; (2) returns from years of surplus schooling are positive, but smaller than returns from years of required education; (3) returns from years of deficit schooling are negative but always smaller in absolute value than the returns from required education; and (4) these findings are robust to different methods of measuring the required education for an occupation, including job content analysis, worker self-assessment and realised matches.

Alternative Model: Verdugo and Verdugo (1989)

In the development, there are some modifications to Duncan and Hoffman's model. For instance, Verdugo and Verdugo (1989) modify the ORU equation by replacing years of deficit schooling and years of surplus schooling by dummy variables of undereducation and overeducation. The model is as follows:

$$\ln W_{i,t} = \beta_0 + \beta_1\, dummy\_X_{i,t}^{\mathrm{UE}} + \beta_2\, dummy\_X_{i,t}^{\mathrm{OE}} + \beta_3\, X_{i,t,n} + \sum_{k=1}^{Q} \beta_{3+k} \mathrm{K}_{i,n,t} + \varepsilon_{i,t} \quad (5.2)$$

where:

---

[57] whereby workers' wages are determined in part by the jobs they are doing.

$dummy\_X_{i,t}^{UE}$: dummy for undereducated; 1 if the n^{th} individual is undereducated, and 0 otherwise; $dummy\_X_{i,t}^{OE}$: dummy for overeducated; 1 if the n^{th} individual is overeducated, and 0 otherwise; and X_n is t
he years of actual completed years of schooling variable. If β_1 is positive and significant, it is interpreted as undereducated workers gaining higher wage (premium) relative to workers with actual completed years of schooling. If β_2 is negative and significant, this implies that overeducated workers receive lower wages than workers with actual completed years of schooling.

Verdugo and Verdugo (1989) use the data from the 1980 US census. In contrast to Duncan and Hoffman (1981) and the human capital theory's prediction, Verdugo and Verdugo find that the coefficient of overeducation is negative whereas the coefficient of undereducation is positive after controlling for education attainment. This suggests that overeducated workers earn less than those who are not overeducated.

There are several possible reasons for this result: (1) the lower return to overeducated workers may reflect their employment in low-paying jobs; and overeducated workers will be concentrated in occupations where workers have low average education, for example elevator operators with one year of college, (2) overeducated workers are unproductive and education does not guarantee an increased productivity, and (3) undereducated workers appear to have an earnings advantage relative to their overeducated (and matched) peers, as they may tend to be excellent performers on their jobs.

Countering Verdugo and Verdugo's (1989) finding, Cohn and Kahn (1995) explore both models by using the same data. The first model is adapted from Sicherman (1991) or similar to Duncan and Hoffman's model, and the second model is from Verdugo and Verdugo. Cohn and Kahn find that a negative estimate on the overeducation dummy for both models does not necessarily imply a negative return to overeducation (wage penalty), that is as long as the coefficient of years of overeducation is significantly positive in the first model. The negative estimate of the years of overeducation coefficient in Verdugo and Verdugo model suggests that overeducated workers earn less than matched workers with similar levels of schooling. The study argues that the result should not be surprising nor conflicting with the finding that the return to overeducation is positive though generally less than the return to matches.

Control Variables

In terms of control variables, the ORU model is the extended model of Mincer wage equation. Thus, the control variables that are commonly used have already been discussed in Chapter 3.

When applying Mincer wage equation, most research modifies the model with control variables, *i.e.* personal characteristics such as sex (Comola and de Mello, 2009), and marital status (Chevalier *et al.*, 2002; Comola and de Mello, 2009); job and firm related variables such as present labour market experience or tenure (Purnastuti *et al.*, 2013), firm size and firm age (Pereira and Martin, 2001), industries (Comola and de Mello, 2013), formal and informal sectors (Dasgupta *et al.*, 2015); urban and rural area residentials (Dumauli, 2015); and some interaction terms such as gender-marital status and gender-dependency ratio (Comola and de Mello, 2009); with the aim of capturing other factors that may affect the wage equation.

Selected studies which used the ORU model and the control variables are shown in Table 5.1. There are some additional variables in these studies, such as: unemployment rate and occupation (Verdugo and Verdugo, 1989) and disability (Tsai, 2000; Dockery and Miller, 2012; and Iriondo and Perez-Amaral, 2013).

5.2.3 Empirical Studies on Mismatch and Wage

There have been many studies which look into education mismatch and wage in developed and developing countries, including Indonesia. However, most studies in Indonesia only review overeducation in the university level. This part will discuss previous empirical studies which use Duncan and Hoffman's model on both overeducation and undereducation. The discussion is structured by methods used, cross-section and panel data. Every sub-section is structured by developed and developing countries. The summary of previous empirical studies is presented in Table 5.1.

Table 5.1: Return to Undereducation, Overeducation and Required Education in Selected Countries (in per cent)

Researchers (Year)	**Country**	**Data Source**	**Year**	**Method**	**Mismatch Variable**	**Return to Under-education**	**Return to Required (Completed) Schooling**	**Return to Over-education**
Cross-Section Studies								
Duncan and Hoffman (1981)	US	The Panel Study of Income Dynamics	1976	OLS: white men	Years of surplus and deficit schooling	-0.42*	0.63*	0.29*
Verdugo and Verdugo (1989)	US	Census	1980	OLS with controls: experience, unemployment rate, region, sector, occupation, marital status, number of weeks unemployed, and number of hours worked.	Dummy variables of UE and OE	0.09**	0.72**	-0.13**
Groot and Van Den Brink (1997)	UK	The British Household Panel Survey	1991	Duncan and Hoffman (OLS)	Years of surplus and deficit schooling	The coefficient values are unspecified, the pattern is similar to Duncan and Hoffman (1981) finding, IV method is required to deal with ability bias.		
Daly *et al.* (2000)	US and Germany	The Panel Study of Income Dynamics	1976, 1985, 1984	Duncan and Hoffman (OLS)	Years of surplus and deficit schooling	US, men, 1976: -3.4** Germany, men, 1984: -7.8**	US, men, 1976: 6.1** Germany, men, 1984: 9**	US, men, 1976: 4.5** Germany, men, 1984: 4.9**
McGuinness (2006)	The US, Canada, Hong Kong, The UK and 6		1980s to 2000s	Meta-analysis		General conclusion: similar pattern with Duncan and Hoffman (1981) finding.		

	EU countries							
Sharma and Sharma (2013)	India	Primary survey	2011-2012	Duncan and Hoffman (OLS)	Years of surplus and deficit schooling	-0.146**		0.103***
Alisjahbana *et al.* (2017)	Indonesia	Labour force survey	2014	Verdugo and Verdugo model (OLS)	Mismatch category	-0.042		-0.321***
Panel Data Studies								
Korpi and Tahlin (2007)	Sweden	The Swedish Level of Living Surveys	1974, 1981, 1991 and 2000	Duncan and Hoffman (OLS)	All education level	-0.03***	0.07***	0.03***
				Duncan and Hoffman (OLS with additional human capital control)	Years of surplus and deficit schooling.	-0.02***	0.06***	0.02***
				Duncan and Hoffman (Fixed Effects/FE)		-0.02***	0.03***	0.01***
Tsai (2010)	US	The Panel Study of Income Dynamics	1979–2005	Pooled OLS: Duncan and Hoffman with controls: sex, age, race, marital status, number of children, disability, tenure, experience.	Dummy variables of UE and OE	0.06***	0.11***	-0.04***
				Random Effect (RE): Duncan and Hoffman with controls: sex, age, race, marital status, number of children, disability, tenure, experience.		0.03***	0.09***	-0.02***
				FE: Duncan and Hoffman with controls: age, marital status, number of children, disability, tenure, experience.		0.00***	0.02***	-0.01**

Dockery and Miller (2012)	Australia	HILDA Survey	2001-2008	Pooled OLS: Duncan and Hoffman with controls: sex, age group, marital status, disability, job status, English competence, work experience.	University level (education level).	-0.04***	0.12***	0.05***
				RE: Duncan and Hoffman with controls: sex, age group, marital status, disability, job status, English competence, work experience.	Years of surplus and deficit schooling.	-0.06***	0.10***	0.05***
				FE: Duncan and Hoffman with controls: sex, marital status, disability, job status, English competence, work experience.		-0.04***	0.06***	0.03***
Iriondo and Perez-Amaral (2013)	11 EU countries	European Union Statistics on Income and Living Conditions (Eurostat)	2006-2009	Pooled OLS: Verdugo and Verdugo, controls: sex, experience, disable and marital status.	Dummy variable of overeducation and undereducation for VV.	0.09***	0.12***	-0.08***
				Pooled OLS: Duncan and Hoffman, controls: sex, experience, disability and marital status.	Years of surplus and deficit schooling for Duncan and Hoffman.	-0.03***	0.14***	0.04***

Yin (2016)	China	The China Health and Nutrition Survey (CHNS)	1989, 1991, 1997, 2000, 2004, 2006 and 2009	Pooled OLS: Duncan and Hoffman with controls: sex, urban, sector, firm size, occupation, job status, experience, region.	All education levels Years of surplus and deficit schooling.	-0.04***	0.23***	0.03***
				Pooled OLS with time effect: ORU with controls: sex, urban, sector, firm size, occupation, job status, experience, region and year		-0.02***	0.03***	0.03***
				FE: Duncan and Hoffman with controls: sector, firm size, occupation, job status, experience and year.		-0.01	0.02*	0.02***
				RE: Duncan and Hoffman with controls: sex, urban, sector, firm size, occupation, job status, experience, region and year.		-0.02***	0.03***	0.03***

Source: The author’s compilation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Studies with Cross-Section Method

The first part is cross-sectional studies in developed economies. Duncan and Hoffman (1981) use the US data (the PSID in 1976)[58] and separate their analysis by sex and race. There are four models: models of white men, black men, white women, and black women. The dependent variable is log of hourly wages while the control variables used are experience and experience-squared, city-size, and a dummy variable for residence in the south (location). All variables are the same for those four models. Duncan and Hoffman find that the surplus of schooling coefficient ($\beta_{2,t}$) for white men was positive (2.9 per cent) and the deficit of schooling coefficient ($\beta_{1,t}$) was negative (-4.2 per cent); all was significant at 1 per cent level. Those values were lower compared to the wages of white men with required education (6.3 per cent). Those values were also relatively smaller than the estimation results for black men. For white women, the coefficient of required education, surplus education and deficit education were 9.1 per cent, 5.2 per cent and -1.4 per cent, respectively. Wage premium for surplus educated workers and wage penalty for deficit educated workers occur due to the productivity being more variable in the US, as explained in Section 5.2.2. From the workers' point of view, it also seems that workers with required education receive higher return than those undereducated or overeducated

By applying equation 5.2 and using data from the US 1980 census (as explained in Section 5.2.2), Verdugo and Verdugo (1989) find that the coefficient of overeducation was -13 per cent and the coefficient of undereducation was 9 per cent after controlling the education attainment. This implies that overeducated workers earn lower relative to workers with actual years of schooling. The coefficient of actual years of schooling was 7.2 per cent.

In the UK, Groot and Brink (1997) investigate Duncan and Hoffman's model by using the 1991 wave of the British Household Panel Survey and by applying the OLS. Groot and Brink identify a pattern similar to that of Duncan and Hoffman's (1981): the highest return is for matched workers, return for overeducated is less than matched, and return for undereducated is negative. Nevertheless, they argue that the result could be overestimated due to ability bias. IV method is therefore required to deal with the bias.

[58] Education is divided into 0-5 grades, 6-8 grades, 9-11 grades, 12 grades, some college, college degree, and an advanced degree. Occupation is divided into professional, manager, self-employed, clerical, sales, craftsmen and foremen, operatives, labourers, service workers and farmers. Meanwhile, mismatch is measured by objective/RM (mean).

Furthermore, Daly *et al.* (2000) evaluate the model by comparing the results based on empirical data in the US and Germany. Germany is chosen considering the country's more structured educational system and labour market. The data used are from the PSID (1976 and 1985 waves) and The German Socio-Economic Panel (1984 wave). Although the data are longitudinal, the analysis conducted is cross-sectional for each year and the Chow test is used for the differences in parameters across the sample. Moreover, the model is modified by adding years of work experience and city residences as control variables. The analysis is then separated by gender. Daly *et al.* find that the pattern and coefficient values (returns) of men for required, surplus and deficit schooling in 1976[59] were 6.1, 4.5 and -3.4 per cent, respectively, which is approaching Duncan and Hoffman's (1981) findings. In 1985, those coefficients slightly increased but the pattern was still the same. For Germany in 1984, the coefficient values (returns) of men for required, surplus and deficit schooling were 9, 4.9 and -7.8 per cent, respectively. In short, this finding is consistent with a universalistic view of labour markets; there are more similarities across countries than over time.

McGuinness (2006) documents mismatch studies (meta-analysis) in the US, Canada, Hong Kong, the UK and six European countries (Germany, The Netherlands, Spain, Portugal, Greece, and Northern Ireland) from 1980s to 2000s periods. McGuinness asserts that the evidence on the return to mismatch is in line with the assignment theory, given that the lower returns to surplus education suggest that overeducated workers work below their potential but are deriving some benefits from the surplus education. These lower returns are consistent with a scenario in which overeducated workers' jobs are imposing an upper limit on the extent to which they can utilise their education and skills with this productivity ceiling reflected in lower wages. The evidence is certainly not consistent with the human capital theory (HCT) which suggests that the returns to surplus and required education should be equal and that the overeducated should not have any pay penalty (lower wages) inflicted upon them, neither is it consistent with Thurow's Job Competition Model which suggests that the return to surplus education is zero.

Turning to developing countries, Sharma and Sharma (2013) study the same issue in the Indian labour market by using the data from the primary survey by the NSSO in 2011-2012 and by employing Duncan and Hoffman's model. The study categorises occupations

---

[59] This is exactly the same data as used by Duncan and Hoffman (1981).

into six different categories: professional, administration, clerical, services, production and farmers. The study observes that the proportion of workers with the required level of education is highest for white-collar jobs and lowest for the blue-collar jobs. This is indicative of insufficient employment opportunities available for the highly educated workers in the Indian labour market. In terms of wages, the results suggest that the return to each surplus year of education (overeducation) is positive and statistically significant, whereas the return to each deficit year of education (undereducation) is negative and statistically significant, similar to Duncan and Hoffman's findings. However, the wage penalty associated with undereducation is higher than the wage gain associated with overeducation.

In Indonesia, this area of research is not well-explored and most existing research focus on higher education levels, as explained in Section 5.1. For instance, Alisjahbana *et al.* (2017) study education mismatch and its effect on the Indonesian labour market by using SAKERNAS data of 2014. The mismatch is estimated by 3-digit occupation level and by employing a modified Verdugo and Verdugo's model. Alisjahbana *et al.* specifically focus on university and vocational diploma in science and engineering graduates who work in the waged sector. The study find that lower wages (wage penalty) exist in the Indonesian labour market, where both the overeducated and undereducated receive lower wages compared to those with adequate educational levels for the job that they hold. The result is in contrast to Duncan and Hoffman model, since Alisjahbana *et al.* only analyse the data based on mismatch status, either overeducated or undereducated, not in terms of extra year of deficit/surplus schooling. Moreover, the effect is greater for university graduates compared to vocational diploma graduates, holding individual and employment characteristics as well as major of study within the science and engineering field and the sector of employment constant. This shows the differential effect of education mismatch between university graduates *versus* vocational diploma graduates, where the former has less effect on its average wage rates.

Studies with Panel Data Method

With regards to studies which apply panel data in developed countries, Korpi and Tahlin (2007) investigate the impact of education mismatch on wages in Sweden by using the level of living surveys from 1974 to 2000. They find that the result replicates the results found in other countries, *i.e.* in the pooled OLS version; the effect of required schooling is positive, the effect of overeducation is also positive although smaller than the effect of required schooling, while the effect of undereducation is negative and smaller in absolute size than the effect of required schooling. Both overeducation and undereducation are significantly different from required schooling estimate. Furthermore, the findings do not support their hypothesis that: (1) education mismatch reflects human capital compensation rather than real mismatch, and (2) education mismatch is real but dissolves with time spent in the labour market so that its impact on wages tends to move toward zero over a typical worker's career.

In the UK, Lindley and McIntosh (2009) examine the impact of overeducation by using the BHPS of the 1991-2005 periods. Pooled OLS and FE are used to allow controlling for unobserved individual heterogeneity in the determinants of incidence and impact of overeducation. ORU is modified by replacing years of schooling with dummy variables. Thus, overeducation is measured by using 5 binary dummy variables and a categorical variable is included to indicate undereducation. Although the method is slightly different, the conclusion is similar to previous empirical studies returns to required education are almost always greater than that to all overeducation levels and the returns for undereducation are negative. Unobserved heterogeneity cannot be the only explanation for the existence of overeducation, though the fact that it has some roles to play is shown by the overeducation penalties being smaller once unobserved heterogeneity is controlled. As a result, the coefficients of FE are lower than the coefficient of pooled OLS.

Tsai (2010) studies the return to education mismatch in the US by using data from the PSID for the period of 1979–2005. The mismatch is measured subjectively. The OLS results reveal that overeducated workers earn significantly less and undereducated workers earn substantially more than matched workers. Using RE and under the assumption that individual-specific error is uncorrelated with the education mismatched variables, the result is similar to the OLS, but with slightly lower coefficients for undereducation and overeducation, presumably because there is an adjustment for the serially correlated components in the error term.

In Australia, Dockery and Miller (2012) investigate the education mismatches and credentialism by using 8 waves of The HILDA survey for the 2001-2008 periods. The study identifies larger returns from years of required education and modest returns from years of overeducation. Workers benefit from being employed in an occupation for which they are undereducated because the positive effect of being in an occupation with a higher reference level of education outweighs the negative effect of their years of undereducation. The most important implications or key policy messages are (1) the additional years of schooling associated with credentialism are not wasted: these additional years appear to be linked to the development of skills that attracts a reward of around 3-6 per cent. This is comforting for advocates of the expansion of the education sector. Also, (2) there are large gains that could be potentially achieved through a better matching of workers' actual educational attainment to their job requirements.

Iriondo and Perez-Amaral (2013) study the effect of education mismatch on wage by using a rich panel dataset of workers in 11 European countries[60] from 2006 to 2009, drawn from the European Union Statistics (Eurostat) on Income and Living Conditions. Pooled OLS, RE, FE and IV-FE are used as comparison. The estimation from Duncan and Hoffman's model using the mean index, the results of the pooled model, and the random-effects model are very similar. In the pooled OLS, wage rises by 14.3 per cent for each year of required education, 4.2 per cent for each year of overeducation, but decreases by 3 per cent for each year of undereducation. When applying FE, the size of the effect falls: wage of required year of schooling becomes 2.7 per cent, and a penalty of 0.5 per cent per year of undereducation occurs. Meanwhile the overeducation coefficient is not statistically different from zero. Iriondo and Perez-Amaral argue that wage basically depends on the educational requirements of the jobs. Verdugo and Verdugo model's results show that the inclusion of overeducation and undereducation incidents increases the return to schooling from 8.0 per cent to 12.0 per cent in the pooled model. Overeducated workers suffer a 7.7 per cent penalty and undereducated workers get a return of 9.3 per cent. RE barely change compared to those obtained in the pooled model. In contrast, FE estimation shows that the size of the coefficients considerably drops to 1.7 per cent in return for each year of schooling, -1.5 per cent for overeducation and 1.0 per cent undereducation.

Turning to developing countries, Yin (2016) estimates the overeducation and wage penalty in China by using the ORU model (OE and UE are dummy variables while required years of schooling is defined as the years of actual completed years of education) with the panel data method and by comparing the mean and the mode. The data used are from the CHNS from 1989, every two to four years and followed by 1991, 1993, 1997,

---

[60] Austria, Belgium, Spain, Finland, France, Greece, Ireland, Italy, Luxembourg, Portugal and the Netherlands.

2000, 2004, 2006 and 2009 (8 waves in total). Yin finds that for the mean method, the wage return to an additional year of required education was around 23 per cent, the wage returns to one year of surplus schooling (overeducation) was around 3 per cent and the coefficient of deficit years of schooling was around 4 per cent. For the mode, the wage returns to required schooling level was 26.20 per cent, the wage return to overeducation was 9.68 per cent and the negative wage return to each additional level of deficit education was about -13.19 per cent.

In other developing countries, Reis (2015) studies wage and mismatch in Brazil. The method used is Duncan and Hoffman's model (1981) and the data are from the Brazilian Census Bureau in the six main Brazilian metropolitan areas between January 2004 and December 2012. This study yields a similar finding and also proves Duncan and Hoffman's results. The coefficient for years of deficit schooling was equal to −0.115, whereas the number of years of surplus schooling was associated with a positive coefficient equal to 0.107. The estimated coefficient for required schooling was 0.170. When taking individuals' fixed effects into account, the estimated impact for each year of surplus schooling on labour earnings dropped to 0.009, whereas the estimated penalty for each additional year of deficit schooling dropped to 0.110.

In short, the literature reviewed above shows that most studies focus their analysis on the relationship between education mismatch and wage in developed nations such as the US, the UK and EU countries. Other studies prove that education mismatch also occurs in developing countries and this incidence affects wage. Thus, studies in this area of interest are required to consider the different characteristics between developed and developing countries. Both Duncan and Hoffman's and Verdugo and Verdugo's model confirm that there is a distinction between wages of the overeducated, undereducated, and matched/adequately educated workers. Using either cross-sectional or panel method, or even using a different data source, all the analyses provide a similar wage pattern of surplus and deficit years schooling for Duncan and Hoffman's model. However, some studies find that unobserved individual heterogeneity may occur in the cross-sectional method, though the panel data model can be used to deal with such issues. In terms of control variables, some studies separate their analysis by gender and race. There is also a potential difference not only by gender, but also by sector, as per Chapter 3 finding. Particularly for Indonesia, Alisjahbana *et al.* (2017) study this issue using cross-sectional method and Verdugo and Verdugo's model. As such, this chapter will use a different model and method (Duncan and Hoffman's and panel method which can deal with heterogeneity issue) in order to contribute to and enrich the existing literature.

5.2.4 The Dynamic of Wage and Mismatch

Some studies explore the dynamic of mismatch in terms of wage persistence for mismatched graduates, whereas some others explore education mismatch, wage growth and the trend of wage penalty overtime.

In terms of persistence, Dolton and Vignoles (2000) examine the effects of overeducation on wage in the UK after graduation and six years later by using Duncan and Hoffman's model. The data are obtained from the UK graduates surveyed in 1980 and 1986. The study finds that 38 per cent of graduates were overeducated for their first job and even six years later. In terms of wage, overeducated graduates earned lower wages, and the wages declined overtime for the same workers. Similarly, Frenette (2004) analyses overeducation among young post-secondary graduates in full-time employment in Canada, particularly two and three years after their graduation. Using the discrete analogue proposed by Verdugo and Verdugo (1989) and the data from the National Graduates Surveys (NGS)[61], the study finds a slight decline in overeducation and wage overtime which occurs once observed heterogeneity is addressed.

Lindley and McIntosh (2010) investigate the impact and permanence of wage for overeducation while controlling unobserved individual heterogeneity. The data used are the first 15 waves (1991-2005) of The BHPS. By using ORU equations with OLS and FE model as the method, the study find that: (1) holding constant job requirements at a particular level will result in variations in the wage penalty suffered by overeducated employees working at that level. In addition, the wage was systematically smaller (only 4 per cent) amongst those workers who would be in a matched job five years later, in 1996. The systematically lower wages at the same job level amongst those with a history of prior overeducation are again taken as an indicator of lower unobserved ability amongst this group. Also, (2) overeducated workers in 1996 and 2001 suffered no wage penalty to their prior overeducation once they are in a matched job. Temporary overeducation

[61] The cohorts are the classes of 1982 (interviewed in 1984/1987), 1986 (interviewed in 1988/1991), and 1990 (interviewed in 1992/1995).

amongst graduates therefore need not be an indicator of lower ability but may simply be a part of the normal work history of some graduates as they acquire the work experience necessary to complement their higher-level qualifications before they can move into graduate-level jobs.

Turning to wage growth, Korpi and Tahlin (2007) analyse education mismatch and wage growth in Sweden by using cross-sectional and panel data from the level of living surveys in 1974-2000. The model used focuses on the impact of being mismatched or matched on all forms of wage growth, in connection to career promotion and otherwise. Korpi and Tahlin argue that overeducated workers should experience greater-than-average wage growth because of several considerations: overeducation may be seen, in a career perspective, as part of a human capital investment strategy; the gap in returns to schooling relative to correctly matched workers should decrease over time and eventually becomes zero; and finally, the mismatch is temporary, not in the planned career sense discussed above but rather as a result of job search with imperfect information. Korpi and Tahlin (2007) also find that wage growth among the overeducated is now significantly lower than among the matched workers. Thus, there is no indication of greater wage growth associated with overeducation. The study concludes that the overeducated are penalised early on by an inferior rate of return to schooling from which they do not recover.

And finally, only a few studies examine the change of wage penalty overtime, such as Green and Henseke (2016). The study analyses the trend of wage penalty in Britain during 1997/2001 and 2006/2012 periods by using the BHPS. Based on the data, the proportion of mismatched graduates in the labour market has remained fairly stable at approximately three-in-ten graduates. In terms of the trend of wage penalty overtime, the study finds that the log wage penalty of mismatched (*i.e.* overeducated) graduates relative to matched graduates was significant and sizable, being 38 log points in 1997–2001 (consistent with other studies). Furthermore, the log increased to 49 points in 2006–2012. Thus, the penalty increased overall by 11 log points. To compute that result, the study runs regressions comprising the full sample of graduates and non-graduates. The result indicates that the premium for matched graduates over matched non-graduates increased significantly by 5 points, which is arguably consistent with the persistence of skill-biased technological change over this period, while at the same time, there was a 6-point significant fall in the premium for mismatched graduates.

5.2.5 Wage and Education Mismatches Based on Gender

In terms of gender, the present study (Chapter 4) finds that male workers tend to have lower education levels than female workers, particularly for at least 13 years of education or equivalent to university level. However, the return to education for females is higher than for males. This could be due to the fact that females have a combination of a much lower workforce participation and fewer average years in the labour force than males (Abbas and Foreman-Peck, 2008); education has increased the female's skills and productivity and has provided an extra return for females (Dougherty, 2005).

In terms of the mismatch trend, males tend to have a higher proportion of overeducation compared to females, while females have a slightly higher proportion of undereducation (Table 4.6). Furthermore, most literature analyses gender effects on mismatch by using a dummy variable in the ORU model and only a few studies separate them (see *e.g.* Duncan and Hoffman, 1981; Mcguinness *et al*., 2010; and Iriondo and Perez-Amaral, 2013). As such, exploring the effect of gender on wage is required to contribute to the existing literature given that the results may differ between sexes.

Firstly, Duncan and Hoffman (1981) extend their analysis based on race and gender (white men, black men, white women, and black women). They find that overeducation has a positive and significant effect on wage rates for all four race-sex sub-groups. Specifically, overeducated white and black women receive substantial lower wages than overeducated men. Undereducated men earn less than the others, including those with the same required level of education, whereas the wages of undereducated white and black women are insignificantly different from those with the same required level of education.

In the UK, Groot and Brink (1997) investigate this aspect by using the 1991 wave of the BHPS and by applying the OLS. They find that the rates of return to a year of education attained were 4.9 per cent for males and 5.2 per cent for females. The rates of return to a year of education required were much higher: 7.7 per cent for males and 8.9 per cent for females. The rates of return to a year of overeducation were also negative: -2.4 per cent for males and -3.3 per cent for females. Meanwhile, the rates of return to one year of undereducation were 5.9 per cent for males and 4.8 per cent for females. In short, males generally received slightly higher returns to required, over and undereducation compared to females Britain in 1991.

McGuinness *et al*. (2010) use the first seven waves of the HILDA survey, starting from 2001. The study finds that controlling unobserved heterogeneity (using RE with Mundlak corrections and FE) could remove most of the wage impacts only on men who are overeducated. Graduate men who change status from a well-matched job to an overeducated-only job do not suffer a wage penalty (lower wages)[62]. In contrast, women tend to have lower wages.

Slightly differently, Iriondo and Perez-Amaral (2013) analyse the IV-FE methods for Duncan and Hoffman's model by gender. The main finding from Duncan and Hoffman's model is that the females in selected European countries have a higher return on attained and required schooling than the males. Nonetheless, wage penalty for one year of schooling has the same pattern as the return to one year of attained schooling in both estimations: the returns to one year of schooling for the overeducated and matched workers are not significantly different from zero for both males and females. This could imply that wage depends mainly on the educational requirements of the jobs and undereducated males tend to suffer a greater penalty in their wages than the females.

5.2.6 Wage and Education Mismatch Based on Sector

A sectoral analysis is also required to have a more comprehensive picture of the Indonesian labour market. The public and private sectors have similar characteristics (receiving regular wages); yet, the public sector is more attractive for job seekers. For instance, the ratio of job opportunities in the public sector to the number of applicants is around 1:200 (Sindo, 2013), as explained in Chapter 4, presumably due to the relative wage in public sector being higher than the wage in the private sector (Chapter 3's finding). In terms of the changes in mismatch, the present study (Chapter 4) also identifies the different trends between the public and private sectors, *i.e.* there is a decrease in overeducation and an increase in undereducation in general and in the private sector in particular. In contrast, the public sector experiences an increase in overeducation and a

[62] Besides studying education mismatch, the study also analyses over-skilled and both overeducation and over-skilled variables. Moreover, it examines the relationship between mismatch and wage, as well as job satisfaction and job mobility. However, the study finds that overeducation and over-skilling are distinct phenomena with different market labour outcomes. The discussion in this part is limited on overeducation only.

decrease in undereducation. Thus, sectors may contribute to the difference in wages between overeducated and undereducated workers in Indonesia. Similar to gender analysis, most of the existing research controls the sector (or occupation) as dummy variables.

In terms of wage, overeducated graduates in the public sector is hypothesised to be less fully utilised and to earn less than the overeducated ones in the private sector due to the relatively less competitive nature of the public sector (Dolton and Vignoles, 2000). Dolton and Vignoles study overeducation in the UK graduate labour market by using a one-in-six sample of the 1980 UK graduates surveyed in 1986. The study limits the analysis by analysing overeducation in the public sector only. The result shows that public sector graduates, in jobs requiring no qualifications, earned 9–10 per cent less than the private sector equivalence. There are several possible explanations for this: more than a third of public sector graduates who were required no qualifications at all were working in welfare jobs. Welfare jobs in the public sector may pay poorly relative to the kinds of jobs requiring no qualifications in the private sector (labouring, sales, industrial processing, *etc*.). Those working in the community and social services sector tend to have the longest job tenure (Allen, 2016), which implies greater security of tenure in the public sector. Moreover, the public pension program is better than the private pension program in the past; the government implemented the public pension programs in 2016, which is based on the government regulation number 45/2015 on the management of the old age security program. Employees in the public and private sectors are covered by both an earning-related social insurance scheme and a defined contribution plan (OECD, 2017). Another possible human capital explanation is that the private sector responds more flexibly; allowing the overeducated graduates to be more fully utilised and paying a wage closer to his or her productive potential.

Different from Dolton and Vignoles (2000), Bauer (2002) does not separate the analysis based on sectors. Instead, all sectors are run in general and another model is added to exclude all individuals working in the public sector. Nonetheless, excluding the public sector does not change the main conclusions; overeducated workers still earn less, and undereducated workers earn more than workers with the same level of educational attainment but who work in occupations that fully utilise their education.

Furthermore, Allen *et al*. (2013) use data from twenty countries[63] to analyse the return associated with education mismatch due to heterogeneous skills or institutional effects. In terms of heterogeneous skills, the reason overeducated workers earn a lower wage is because they have a lower level of skills and are consequently sorted into in lower-level jobs, compared to their more skilled peers who find jobs at their own levels.

Turning to institutional effect, wages are often not based directly on workers' productivity in the job but come about as a result of a process of bargaining involving specific wage-setting institutions. Allen *et al*. (2013) conclude that wage-setting institution theory explains observed wage effects in the public sector, while heterogeneous skill theory explains the one in the private sector. Particularly for the public sector, the study finds that overeducated workers are likely to be paid higher wages than one would expect based on their productivity. This view is further warranted by the finding that the private sector wage's effects on overeducation are especially strong in countries with high levels of relative heterogeneity. Moreover, there is a strongly significant negative interaction between collective bargaining coverage and the effect of overeducation, which suggests that lower wage (stronger wage penalty) in the public sector is stronger in countries where there is a high level of collective bargaining coverage than in countries with lower levels of collective bargaining coverage (Allen *et al*., 2013).

5.3 Model, Method and Data

5.3.1 Model

Considering data availability as well as the advantages and disadvantages of the model, the present study applies the Duncan and Hoffman's model as it can analyse and compare the return to undereducation and overeducation, in terms of additional deficit or surplus years of schooling, which is consistent with the aims and research questions of this research. Besides, the model also offers some advantages, such as better interpretation of

---

[63] Austria, Belgium, the Czech Republic, Estonia, Finland, France, Germany, Hungary, Italy, Japan, Lithuania, the Netherlands, Norway, Poland, Portugal, Slovenia, Spain, Switzerland, Turkey and the United Kingdom. The sample was graduates who obtained a higher education degree in the 1999/2000 academic year.

mismatch-wage relationship than the Verdugo and Verdugo's model[64], and the most important thing, the model has proven itself a reliable extension of Mincer wage equation by passing statistical testing in several countries for several periods based on previous empirical studies, as explained in the literature review (some of them are provided in Table 5.1).

The models are as follows:

$$\ln W_{i,t} = \beta_0 + \beta_1 X^{\mathrm{UE}}_{i,n,t} + \beta_2 X^{\mathrm{OE}}_{i,n,t} + \beta_3 X^{\mathrm{REQ}}_{i,n,t} + \sum_{k=1}^{Q} \beta_{3+n} \mathrm{K}_{i,n,t} + \varepsilon_{i,t} \qquad (5.3),$$

$\ln W_{i,t}$: log of real hourly wage (similar dependent variable to Chapter 3, section 3.4.1).

In line with Duncan and Hoffman, the model requires variables: (1) years of required schooling ($X^{REQ}_{i,n,t}$), which is equal to the mode of each hybrid occupation category, as explained in Chapter 4, Section 4.3.2; (2) years of deficit schooling relative to the mode level of education in the individual's occupational category (refers to undereducation or $X^{UE}_{i,n,t}$), or equal to $X^{REQ}_{i,n,t} - X$, and X is actual years of schooling. $X^{OE}_{i,n,t}$ or years of surplus schooling relative to the mode level of education in the individual's occupational category (refers to overeducation), calculated from $X - X^{REQ}_{i,n,t}$. Years of schooling is a continuous variable and has 7 categories from 6 to 22 years of schooling. Similar to Chapter 4, the objective method (by deviation from mode) and the realised method (RM) are used to determine the mismatch, considering the data availability in those IFLS waves. Furthermore, mode is used since it has better estimates for skewed or non-skewed distributed data and is more suitable for categorical data like match variable. Meanwhile, the hybrid category of occupation (44 occupation categories) is constructed with the aim of reducing heterogeneity and avoiding low number of observations.

The control variables used here are the same as in Chapter 3 (Part 3.3.1); thus, all definitions and formulas of calculation follow Chapter 3. However, considering the finding in Chapter 3, some variables are adjusted for simplification. For instance, provinces are simplified into capital and non-capital regions, given that most companies have their headquarters in the capital regions. Also, ethnicity is eliminated here because those dummy variables are mostly insignificant (see the results in Chapter 3). The control variables in the present study are experience and experience squared, and the other control variables are divided into some categories:

[64] Cohn and Kahn (1995) argue that there may be a misinterpretation of Verdugo and Verdugo's model.

1. personal characteristics; consisting of dummy variables for gender (1=female) and marital status (1=single, 2=married and cohabitate, 3=other status).
2. work related and firm size; consisting of full-time/part-time dummy (1=full-time[65] and 0=part-time), job market experience and its square (in years), tenure or the number of years with the current employer and its square (in years), sectoral dummy (1=private and 2=public), and industry dummies (agriculture; mining and quarrying; manufacturing; electricity, gas and water; construction; wholesale retail restaurant and hotels; transportation storage and communications; finance, insurance, real estate and business services; and social services). Agriculture and other sectors are omitted variables and firm size here refers to the number of workers in the firm[66]; and
3. regional dummy variables; consisting of urban/rural dummy and capital/non-capital province residence.

For the pooled OLS method, the present study adds dummy of wave (time), comprising of dummy for 2000, 2007 and 2014 periods, with the baseline of the 2000 period. This is discussed further in the following part.

5.3.2 Methods: Panel Data with Pooled OLS, Fixed Effect (FE) and Random Effect (RE) Model

The first method is the pooled OLS. Hartog (2000) asserts that the ORU model (Duncan and Hoffman's specification) assumes that unobserved heterogeneity (including ability, motivation, compensating differentials and other unobserved characteristics) is uncorrelated with education mismatch. There is the problem of omitted variable bias if this assumption fails to hold. Furthermore, the omission of unobserved heterogeneity may lead to underestimation of the rate of return to overeducation if unobserved heterogeneity is negatively correlated with overeducation; conversely, the rate of return to

[65] A full-time worker is defined as someone working equal to or more than 30 hours per week; this is based on ILO which uses 30 hours per week as the cut-off point for its definition of a part-time worker (Felipe and Hasan, 2006).

[66] Slightly different from Chapter 3, the dummies of firm size are only of three categories: firms with 0-19 workers (small enterprise), 20-99 workers (medium enterprise) and more than 100 workers (big enterprise). The present study eliminates 0-5 workers as the benchmark, considering some potential bias from comparing very small companies with medium/large companies.

undereducation is probably overestimated in case of a positive correlation between unobserved heterogeneity and undereducation.

In addition, clustered/robust standard errors (CSEs) can be used in regression models where observations can be grouped into clusters, with model errors uncorrelated across the clusters but correlated within the clusters; the cluster used for educational mismatch is usually by individuals, year and/or occupation. This method aims to obtain correct statistical inferences; many empirical applications feature the potential for errors to be correlated within clusters. Failure to control within-cluster error correlation can lead to: (1) standard errors that are smaller than regular OLS standard errors, (2) narrow confidence intervals, (3) T-statistics that are too large, and (4) misleadingly small p-values (Cameron and Miller, 2015). The present study will elaborate this method and estimation that clustered by individual and by year in Section 5.4.1 and Appendix XVII, this is to control individual heterogeneity over time. Using this, the results of the pooled OLS with dummy year and clustered standard errors by individuals are exactly the same.

Furthermore, one of the main advantages of using panel data in this context is dealing with individual heterogeneity issue. Theoretically, panel data are advantageous since they include a much larger dataset, thus providing more information, more variability, less collinearity among the variables, more degrees of freedom and more efficiency. The models are definitely attractive and appealing since they provide ways of dealing with heterogeneity (Park, 2011). On the other hand, panel data may also have some limitations, such as: selectivity problems[67] and cross-sectional dependence (Mućk, 2018).

However, few emerging studies that have applied panel data techniques to analyse overeducation have found controversial results on whether or not education and occupational mismatch with earnings penalty effects are verified once individual heterogeneity is taken into account (Wen and Maani, 2017). For instance, Tsai (2010) finds that the estimated wage effect of overeducation in FE model is very small. In contrast, Dolton and Vignoles (2000) as well as Leuven and Oosterbeek (2011) find the effect is significant; and the coefficient values of education mismatch in the FE model is lower than the coefficient values of the other models. To deal with this, some previous empirical studies attempt to find proxies of unobserved individual heterogeneity, such as health and verbal ability (Korpi and Tahlin, 2006), some of which use panel data model.

[67] Includes self-selectivity, attrition and non-response.

This is because panel data model accommodates individual effect. Thus, panel data can capture features of an individual such as motivation and ability that are given and assumed constant over time (Nielsen, 2014).

Fixed Effect (FE) versus Random Effects (RE)

The crucial distinction between FE and RE is whether the unobserved individual effect embodies elements that are correlated with the regressors in the model, not whether these effects are stochastic or not. The main assumptions of RE model are random and uncorrelated with independent variable included in the model. The advantage of FE model is that it can be used to analyse the impact of variables that vary over time, so the estimated coefficients of FE models cannot be biased because of omitted time-invariant characteristics, such as culture, religion, gender, race, *etc*. Meanwhile, the main limitations of FE depend on the number of characteristics being studied, the degrees of freedom increased, and whether the estimations are linear.

In comparison, the coefficients of FE are usually lower than the coefficients of RE or pooled OLS. The decrease in the estimates when using FE estimation could be caused by measurement error, such as due to required schooling varying within occupation. For instance, the required education for a secretary job in a large professional company might differ from that in a small business. The other possible cause is heterogeneity issue, as Tsai (2010) argues that workers with higher education qualifications have relatively higher ability; thus, the negative relationship between overeducation and ability should be weaker.

FE modelling is used more frequently in economics and political science, reflecting its status as the “gold standard” default (Schurer and Yong, 2012). However, RE models, also called multilevel models, hierarchical linear models, and mixed models, arguably would be the preferred choice because of its greater flexibility and generalisability, as well as its ability to model context, including variables that are only measured at a higher level (Bell and Jones, 2015). The assumptions made by RE models, including the exogeneity of covariates and the normality of residuals, are at least as reasonable as those made by FE models when the model is correctly specified. Moreover, a set of longitudinal panel data usually has the problem of hierarchies[68] occuring in the data, which occur when the population is hierarchically structured. A problem of hierarchies in the data could also be imposed during data collection. Bell and Jones also assert that the downside of RE modelling – correlated lower-level covariates and higher-level residuals – is the omitted-variable bias, which is solvable with Mundlak's (1978) formulation.

Meanwhile, Nielsen (2007) argues that there is another consideration when choosing between FE and RE model, not only by the Hausman test. Unobservable earning ability can be correlated with the mismatch variables, ethnicity or the other explanatory variables since FE uses only within-individual variations. An individual who changes his job to a different occupational category, but without changing his level of education, will possibly have different values of $X_{i,n,t}^{REQ}$, $X_{i,n,t}^{UE}$, and $X_{i,n,t}^{OE}$. By the construction, $X_{i,n,t}^{REQ} + X_{i,n,t}^{OE} - X_{i,n,t}^{UE} = X$, which is constant. As a result, the within-individual variation in $X_{i,n,t}^{REQ}$, $X_{i,n,t}^{UE}$, and $X_{i,n,t}^{OE}$ is characterized by perfect multicollinearity for persons whose education level is constant in the estimation period. Consequently, the FE model would only identify the effect of overeducation or undereducation from information on individuals who change their level of education within the sampled period. Nielsen also contends that very few individuals in the dataset change status from category of workers with required years of schooling to being surplus years of schooling. Thus, RE estimation is preferable, despite its more restrictive assumption that individual effect is uncorrelated with the exogenous variables.

Post Estimation Test

Several post estimation tests are performed in the present study to find out the most appropriate model. Firstly, Breusch-Pagan Lagrange multiplier (LM) test is used to identify the more favourable model between a RE and a simple OLS regression. The null hypothesis (H0) in the LM test is that variances across entities are zero. This means that

[68]With hierarchical data, particularly with temporal hierarchies which are often characterised by marked dependence over time, this is a patently unreasonable assumption. Responses for measurement occasions within a given higher level entity are often related to each other. Subsequently, the effective sample size of such datasets is much smaller than a simple regression would assume, closer to the number of higher-level entities (individuals or countries) than the number of lower-level units (measurement occasions). As such, standard errors will be incorrect if this dependence is not taken into account (Moulton, 1986).

no significant differences occur across units (*i.e.* no panel effect). If the null hypothesis cannot be rejected, this implies that RE is more favourable than pooled OLS.

The next test is the Hausman test, which can be run to decide between FE and RE. The null hypothesis (H0) is that the appropriate model is RE, when there is no correlation between the error term and the independent variables in the panel data model. The alternative hypothesis (H1) is that the appropriate model is FE, when the correlation between the error term and the independent variables in the panel data model is statistically significant. The null hypothesis (H0) is rejected if the Hausman statistic is bigger than its critical value. It is worth noting that RE models have the additional assumption that the individual effects are randomly distributed. It is not just the opposite of FE model, but rather a special case. If the RE assumption holds, then the RE model is more efficient than the FE model (Baltagi, 2015; Bell and Jones, 2015).

The present study also performs the test of equality on two coefficients; between coefficients of years of surplus schooling (OE) and coefficient of years of required schooling (REQ). This assumes that the years of surplus schooling (OE) and the years of required schooling (REQ) are on the same scale. The hypotheses are that the years of required is equal to the years of deficit schooling and that years of required schooling is equal to years of surplus schooling. If those hypotheses are rejected, this implies that the coefficients are statistically different.

5.3.3 Data

Sample and Data Restrictions

The present study uses IFLS data (similar to previous chapter) from 2000 (IFLS3), 2007 (IFLS4) and 2014 (IFLS5). Adding the 2007 data has the benefit of increasing the number of observations. Those waves of survey also represent the condition before the education reform period (2000), during the initial period of the reform (2007) and after the reform (2014), as explained in Section 5.1.

It is worth noting that IFLS has five waves up to 2014. 17,295 respondents took part in all 5 waves or between 1993 and 2014; this is around 52.3 per cent of household members in IFLS1 (1993), of which figure 11,889 (around 54 per cent of the main respondents) had interviews in all five waves. In terms of the re-contact rate, IFLS3 (2000) has 95.3

per cent re-contact rate of IFLS1 (1993) households. Meanwhile, IFLS4 (2007) has 93.6 per cent re-contact rate of the original IFLS1 dynasties and the latest wave (IFLS5 - 2014) has 92 per cent re-contact rate of IFLS1, as explained in Section 2.2.2.

Sample restrictions are also applied on the estimations to ensure that the individuals are in the labour market. Similar to previous chapters, the main data restriction is age (between 16-55 years old) to ensure that the sample is within the age range between finishing compulsory education and retiring. Furthermore, restricting the data based on employment eliminated around 30 per cent of the sample in 2000 and 2007 and around 18 per cent in 2014. This is because some of them are in working age but not in the labour market, for instance, housewives and university students. Moreover, the loss from age restriction to employment status (non-missing) was significantly lower in 2014 than in the other waves (17.87 per cent of loss), possibly because the 2014 wave had casual employment categories which accommodate more workers to fill the questionnaire[69]. For the waged sector (public and private), this research considers the same characteristics, *i.e.* receiving regular wages. This restriction eliminated more than half of the sample in the employment data, which implies that the number of workers in the waged sector was lower than workers in the other sectors. A slight loss also occurred in mismatch (non-missing data) because the sample did not fill the occupation data in their questionnaire. The table of sample restriction is shown in Table 5.2.

It is worth noting that panel data have two approaches in terms of their analysis, *i.e.* balanced and unbalanced panel data. In a balanced panel, each individual (unit) has the same number of observations in all time periods. Since it allows an observation of the same individual (unit) in every period of time, the main advantage of a balanced panel data is minimising the heterogeneity problem. On the other hand, the sample of a balanced panel would be smaller than the sample of an unbalanced panel. The downside of a balanced panel is that sample representativeness will fall over time, particularly if attrition is non-random. Yet, more generally a panel approach does mean that the later waves are skewed towards older age groups. Meanwhile, an unbalanced panel is a panel in which

[69] According to Statistics Indonesia (2010), around 60-70 per cent of the workforce is estimated to engage in informal employment, as explained in Chapter 4.

the number of time series observations is different across units; as long as observations are missing at random[70], an unbalanced panel data are sufficient for the analysis.

From those three waves, the total number of observations become 21,174, consisting of 6,385 observations for the 2000, 6,300 observations for 2007, and 8,489 observations for 2014 (Table 5.2). Therefore, the unbalanced panel are employed in this part of analysis. Meanwhile, the balanced panel data are used in Section 5.4.5 (Robustness test).

Table 5.2: Data Restrictions, 2007

	2000		2007		2014	
	Number of observations	**% Lost**	**Number of observations**	**% Lost**	**Number of observations**	**% Lost**
All individuals	25,825		29,967		36,381	
Age 16-55	21,100	-18.30	25,162	-16.03	29,797	-18.10
Employment status data (non-missing)	14,771	-30.00	17,620	-29.97	24,473	-17.87
Employment status (Government worker, Private worker)	6,780	-54.10	7,148	-59.43	8,712	-64.40
Mismatch (non-missing)	6456	-4.78	6,930	-3.05	8,712	0.00
Wage (non-missing) and other control variables	6385	-1.10	6,300	-9.09	8,498	-2.46

Source: The author's calculation.

Longitudinal Elements of the Survey

Using panel data element, the actual number of individuals reduces to only 15,440 individuals from those observations (21,174 observations), as shown in Table 5.3. Individuals who participated in all those three waves were around 7.3 per cent of the sample (1,128 individuals). 579 individuals in 2014 were interviewed in 2000 (excluding the sample that participated in all three waves). Also, there were 3,477 individuals who were only interviewed in 2000 (discontinued sample). Most of the respondents in the latest wave were new; more than 5,000 individuals for 2014, which follows the rules expansion in IFLS3 in determining the new sample, as explained in Section 2.2.2. In total,

[70] Missing at random (MAR) means that the missing data could be a non-random subset of the data, but that the non-randomness can be completely explained by variables that are in the data.

there are 1,128 individuals who participated in all three waves, and from this sample, a balanced panel data analysis can be conducted. The rows reflect mismatch status in initial year or in the first table is year of 2000, the columns reflect the mismatch status in the later periods. The first row of the table indicates some of 55.73 per cent of undereducated workers in 2000 remained undereducated in 2007. 41.4 per cent of undereducated workers became matched in 2007. And only 2.87 per cent of undereducated workers became overeducated in 2007.

Table 5.3: Number of Individuals and the Transitions Based on Panel Data Element

Number	Pattern	Total		
		Freq.	**Per cent**	**Cum.**
1	2014	5,084	32.93	32.93
2	2000	3,477	22.52	55.45
3	2007	2,273	14.72	70.17
4	2007 and 2014	1,698	11	81.17
5	2000 and 2007	1,201	7.78	88.94
6	2000-2007-2014	1,128	7.31	96.25
7	2000 and 2014	579	3.75	100
		1,5440	100	

Transition: 2000-2007

	2007		
2000	UE	M	OE
UE	175	130	9
	55.73	41.4	2.87
M	317	795	175
	24.63	61.77	13.6
OE	87	317	324
	11.95	43.54	44.51

Transition: 2007-2014

	2014		
2007	UE	M	OE
UE	330	181	30
	61	33.46	5.55
M	245	1,046	278
	15.62	66.67	17.72
OE	60	313	343
	8.38	43.72	47.91

Source: The author’s calculation.

Main Variables Description

The main variables of the present study are wage and mismatch. Firstly, wage refers to the log of hourly wage. The advantage of using hourly wage is to eliminate unobserved heterogeneity caused by the omitted working hours (Li and Urmanbetova, 2007). Hourly wages are equal to last year wages (in real term) divided by the total number of hours per week times the total number of weeks per year, as explained in Section 3.3.

The second variable is mismatch, which is divided into three categories, *i.e.* workers with years of required schooling (similar to match category in Chapter 4), workers with surplus years of schooling and workers with deficit years of schooling[71]. Before discussing the mismatch, Table 5.4 shows the sample distribution based on the actual years of schooling. Most of the respondents have 12 years of actual schooling (36.7 per cent) or equal to senior high school level; followed by 6 and 9 years of schooling. This indicates that most individuals still have low education attainment. Only around 22 per cent of the sample have more than 12 years of actual schooling. There were also 2 observations with 21 years of schooling in 2007, but the sample size was very small, and the proportion was nearly zero per cent.

Table 5.4: Sample Distribution Based on the Actual Years of Schooling, in per cent

Years of Schooling	**2000**	**2007**	**2014**	**Total**
6	36.9	21.1	17.2	24.3
9	18.5	15.6	16.5	16.8
12	31.5	38.1	39.4	36.7
15	6.0	8.8	5.9	6.8
16	7.0	15.5	19.2	14.4
18	0.2	0.8	1.7	1.0
21	0.0	0.0	0.0	0.0
22	0.0	0.0	0.1	0.0
Total Sample	6,385	6,300	8,489	21,174

Source: The author's calculation.

[71] Individuals are considered to have deficit year of schooling if they have actual years of schooling below their occupation's mode. Workers have surplus year of schooling if they have actual year of schooling above their occupation's mode. Required year of schooling is equal to the mode of each occupation category. The difference between mismatch variables in Chapter 4 and Chapter 5 are: Chapter 4 uses education level to estimate the mismatch (for example: 1 for primary school and 2 for junior high school), while Chapter 5 uses year of schooling (primary school represents 6 year of schooling and junior high school represents 9 year of schooling).

Table 5.5 shows the sample distribution based on years of required schooling, which is equal to the mode of each occupation. Thus, the sample proportion of years of actual schooling is different from that of years of required schooling. Most individuals have 12 years of required schooling (around 51 per cent of the total sample), followed by those with 6 years of schooling. Meanwhile, only 1.1 per cent of the sample has a mode of 9 years of schooling. Table 5.5 also indicates the decreasing trend of 6 years of required schooling (low education qualifications) and the increasing trend of 16 years of required schooling or equivalent to university qualifications. Workers with low educational qualifications are needed as production workers in the manufacturing sector (Suryahadi *et al.*, 2003). Meanwhile, higher education qualifications are required considering the rapid technological changes.

Table 5.5: Sample Distribution Based on Years of Required Schooling, in per cent

Years of Required Schooling	**2000**	**2007**	**2014**	**Total**
6	56.1	27.7	22.4	34.1
9	0.9	2.9	0.0	1.1
12	35.9	54.5	60.2	51.2
15	7.1	2.1	2.5	3.8
16	0.0	12.8	14.8	9.8
Total Sample	6,385	6,300	8,489	21,174

Source: The author's calculation.

For mismatched workers, most individuals have years of surplus of schooling around 3 to 6 years from the years of required schooling. 74.6 per cent of the sample who has zero years of surplus schooling is either in the required or deficit years of schooling (Table 5.6). In terms of the trend, around 11.2 per cent of the sample in 2000 had 3 years of surplus schooling. The percentage decreased to 6 per cent in 2014. Similarly, the sample with 6 years of surplus schooling decreased from 11.7 per cent in 2000 to 6.8 per cent in 2014. In contrast, the sample proportion with 2, 4 and 9 years of surplus schooling increased between 2000 and 2014. The majority of the individuals experienced a decreasing trend of years of surplus schooling or overeducation, confirming the finding in Chapter 4 that overeducation proportion decreased between 2000 and 2014.

Table 5.6: Sample Distribution Based on Years of Surplus Schooling, in per cent

Years of Surplus Schooling	2000	2007	2014	Total
0	71.5	75.0	76.6	74.6
2	0.0	0.3	0.8	0.4
3	11.2	7.2	6.1	8.0
4	3.5	5.9	6.8	5.5
5	0.0	0.0	0.0	0.0
6	11.7	8.1	6.8	8.7
7	0.0	0.1	0.0	0.0
9	1.0	0.9	0.8	0.9
10	1.1	2.3	1.9	1.8
12	0.0	0.1	0.2	0.1
16	0.0	0.0	0.0	0.0
Total Sample	6,385	6,300	8,489	21,174

Source: The author's calculation.
Notes: If the years of surplus schooling equals to zero, it means that the individuals are either matched or undereducated.

For years of deficit schooling, most individuals have a deficit schooling of around 3 to 6 years from the years of required schooling (Table 5.7), which is similar to the proportion of years of surplus schooling. The sample proportion with 3 years of deficit schooling increases from 7.2 per cent in 2000 to 10 per cent in 2014. The proportion of workers with 6 years of deficit schooling also slightly increases from 6.3 per cent to 9 per cent in the same period. A similar trend occurs for the other categories of years of deficit schooling as well. This also confirms the findings in Chapter 4 that there is an increase in undereducation.

Table 5.7: Sample Distribution Based on Years of Deficit Schooling, in per cent

Years of Deficit Schooling	2000	2007	2014	Total
0	86.4	79.4	77.9	80.9
3	7.2	8.7	10.0	8.8
4	0.0	2.0	2.3	1.5
6	6.3	9.2	9.0	8.3
7	0.0	0.4	0.4	0.3
9	0.1	0.1	0.0	0.1
10	0.0	0.2	0.3	0.2
Total Sample	6,385	6,300	8,489	21,174

Source: The author's calculation.
Notes: If the years of deficit schooling equals to zero, it means that the individuals are either matched or overeducated.

By gender, the same trend also occurs for both males and females, as most individuals had either 3 or 6 years of deficit schooling compared to required schooling, or around 8 per cent of the total sample (see Appendix XVI). With regards to years of surplus schooling by gender, most males had 3 or 6 years of surplus schooling. Slightly different to males, most females had around 3-5 years of surplus schooling. The trends were similar for all individuals, for both males, and females, *i.e.*: there were a decrease in the sample proportion with years of surplus schooling (overeducation) and an increase in the sample proportion with years of deficit schooling (undereducation).

In terms of sector, the private sector had a similar trend with all individuals: most of the years of deficit and surplus schooling were either 6 or 3 years (see Appendix XVI). However, the public sector had a different pattern: the highest proportion of years of deficit schooling was three years. This is possibly because the public sector workers are expected to have university degree or at least three-years diploma qualifications. Moreover, the highest proportion of surplus schooling was 4 years. In terms of trend, the pattern was different between both sectors. In the private sector, the sample proportion with 3 years of deficit schooling (undereducation) increased from 6.3 per cent in 2000 to 11.2 per cent in 2014. In contrast, the sample proportion of with 6 years of deficit schooling in the public sector decreased from 12 per cent to 4.2 per cent in the same period of time. For overeducation, the sample proportion with 3 years of surplus schooling decreased in the same period of time. However, the sample size in the public sector was smaller than the sample size in the private sector. The overall trend reflects the change in the private sector rather than the public sector. This finding is also consistent with Chapter 4.

Summary Statistics

Turning to the summary statistics (Table 5.8), the mean of log of hourly wage is 8.29, with the minimum 0 and the maximum 16.94. Years of required schooling has a mean value of 10.42 years (equal to senior high school level). The mean of years of surplus of schooling is 1.27 years from the required schooling and the mean of years of deficit schooling (undereducation) is 0.86 years (less than 1 year of schooling). It is worth noting that the other workers are assumed to spend zero year of schooling; thus, the minimum is 0 in the table.

With regards to experience as the first control variable, the present study uses potential experience with a similar formula as in Chapter 3. As a result, 221 observations have negative or zero value of experience; this negative value could arise because the potential experience is formulated as: experience of workers with primary school degree or below is age (in years) – 12 years; experience of workers with junior high school degree is age – 15 years; experience of workers with senior high school degree is age – 18 years; and experience of university qualification is age – 22 years (Dong, 2016), and there is a deviation of school entering age, for example: workers aged 16 years old, he/she already graduated from senior high school, the potential experience is is -2. Then, the negative values are replaced by 0, following ACAPS (2016). This also implies the measurement error in experience could occur for example for individuals who graduate early. The mean of potential experience is 15.62 years. Another control variable is sex. Here, sex is a dummy variable with the mean of 0.37. This implies that the proportion of males is higher than the proportion of females in the present study.

For ethnicity, there are some inconsistent data across the waves. Inconsistent data could result from inconsistent answers from the respondents or a careless error. The present study follows one of the solutions for inconsistent data from Merckle *et al.* (2015) by choosing responses from the 2000 survey wave, *i.e.* the earliest responses. This is based on the assumption that cultural experiences "sediment" as a person grows older, meaning that errors are due to respondents forgetting the answers they supplied in the previous surveys. Therefore, correcting them means accepting the initial response as accurate and the following ones as inaccurate. From the summary statistics, only around 16 per cent of observations are the majority ethnicity.

For marital status, most individuals are married: the mean value of married and cohabitate status is 0.7, which is higher than the other marital statuses. In terms of employment status, most individuals work as full-time workers who work for more than 30 hours per week. Furthermore, some individuals have tenure higher than experience. Thus, tenure has higher maximum values than experience. This is because tenure uses actual value, as it is questioned in the survey, while experience is the potential one, as explained previously. As it is only a small proportion, the present study sets those negative values to zero.

Table 5.8: Summary Statistics of Main Variables and Selected Explanatory Variables

Variable	Obs.	Mean	SD	Min	Max
Wage (log of hourly wages)	21,174	8.29	1.79	0	16.94
Years of required schooling	21,174	10.42	3.44	6	16.00
Years of surplus schooling (OE)	21,174	1.27	2.42	0	16.00
Years of deficit schooling (UE)	21,174	0.86	1.90	0	10.00
Potential Experience	21,174	15.62	10.22	0	43.00
Potential experience squared	21,174	348.56	391.07	0	1849.00
Sex (1=female)	21,174	0.37	0.48	0	1
Ethnicity (1=Javanese)	21,174	0.16	0.36	0	1
Marital status: Single	21,174	0.26	0.44	0	1
Marital status: Married and cohabitate	21,174	0.70	0.46	0	1
Marital status: Other (Separated, divorced and widowed)	21,174	0.04	0.20	0	1
Employment Status: full-time (30 hours a week or more)	21,174	0.84	0.36	0	1
Tenure	21,174	6.49	7.34	0	52.00
Tenure squared	21,174	96.00	196.32	0	2704.00
Sector: private	21,174	0.83	0.38	0	1
Industry1: agriculture	21,174	0.10	0.30	0	1
Industry2: mining and quarrying	21,174	0.01	0.11	0	1
Industry3: manufacturing	21,174	0.22	0.41	0	1
Industry4: electricity, gas and water	21,174	0.01	0.09	0	1
Industry5: construction	21,174	0.06	0.24	0	1
Industry6: wholesale, retail, restaurants and hotels	21,174	0.15	0.36	0	1
Industry7: transportation, storage, and communications	21,174	0.04	0.19	0	1
Industry8: Finance, insurance, real estate and business services	21,174	0.04	0.19	0	1
Industry9: Social services	21,174	0.37	0.48	0	1
Firm size1: 1-19 people	21,174	0.60	0.49	0	1
Firm size2: 20-99 people	21,174	0.24	0.43	0	1
Firm size3: >= 100 people	21,174	0.16	0.37	0	1
Urban	21,174	0.68	0.46	0	1
Capital region	21,174	0.11	0.31	0	1

Source: The author's calculation.

In terms of sector, most individuals work in the private sector, similar to Chapter 3. Related to industry dummies, the three highest means are social services (0.37), manufacturing (0.22), and wholesale, retail, restaurants and hotels (0.15). Furthermore, there are three dummies for firm's size, which represents small, medium and large firms (based on the number of workers). The sample is dominantly small firms with 1-19 workers. And finally, most sample resides in urban areas, around 68 per cent, though only 11 per cent of them reside in the capital province (DKI Jakarta).

Furthermore, Table 5.9 shows that females have slightly higher mean of years of required schooling than males, 10.96 and 10.11 years, respectively. Males seem to have higher years of surplus schooling, around 1.41 years higher than females. In contrast, females have higher years of deficit schooling than males. In terms of wages, males have a slightly higher mean of log of real hourly wages than females, 8.39 and 8.11, respectively. This is in line with Becker's employer taste model of discrimination (Becker, 1971), as explained in Chapter 3, and the statistical discrimination theory.

The values of sector (Table 5.9), wage, years of required schooling, years of surplus schooling (overeducation), experience and tenure in the public sector are relatively higher than in the private sectors. In contrast, years of deficit schooling (undereducation) is relatively higher for the private sector than for the public sector: 0.94 and 0.47 years of undereducation, respectively. This is in line with Part 3.4.3; the public sector in Indonesia prefers workers with high education levels; and wages in the public sector are relatively higher than wages in the private sector. Full table of the summary statistics is presented in Appendix XVI.

Table 5.9: Summary Statistics of Main Variables, by Gender and Sector

By Gender	Male					Female				
	Obs	Mean	SD	Min	Max	Obs	Mean	SD	Min	Max
Wage (log of hourly wages)	13,312	8.39	1.76	0	16.94	7,862	8.11	1.84	0	14.20
Years of required schooling	13,312	10.11	3.33	6	16.00	7,862	10.96	3.54	6	16.00
Years of surplus schooling (OE)	13,312	1.41	2.50	0	12.00	7,862	1.02	2.26	0	16.00
Years of deficit schooling (UE)	13,312	0.82	1.87	0	10.00	7,862	0.93	1.96	0	10.00

By Sector	Private					Public				
	Obs	Mean	SD	Min	Max	Obs	Mean	SD	Min	Max
Wage (log of hourly wages)	17,512	8.12	1.80	0	16.94	3,662	9.08	1.55	0	13.88
Years of required schooling	17,512	9.99	3.24	6	16.00	3,662	12.48	3.58	6	16.00
Years of surplus schooling (OE)	17,512	1.15	2.25	0	16.00	3,662	1.81	3.04	0	16.00
Years of deficit schooling (UE)	17,512	0.94	1.99	0	10.00	3,662	0.47	1.38	0	10.00

Source: The author’s calculation.

5.4 Estimation Results

The following parts will further discuss the estimation results of the ORU model in Indonesia for 2000, 2007 and 2014 periods. The analysis is conducted for the main model (all individuals), based on gender and sector; followed by a brief discussion on control variables’ findings and robustness tests.

5.4.1 Wage Effect: All Individuals (The Main Model)

Table 5.10 reports the estimated results of the ORU model in all individuals. The present study estimates the pooled OLS with the dummy of year, RE, and FE. In addition, pooled OLS model without the dummy of year is provided in Appendix XVII. The interpretation of the coefficient estimates of the control variables are *ceteris paribus*.

Based on pooled OLS, the wage return to an additional year of required schooling was around 10.6 per cent. The wage returns to one year of surplus schooling was around 10 per cent, which was slightly lower than the wage returns to years of required schooling. Meanwhile, the wage returns to one year of deficit schooling was -8.7 per cent. All of the variables were significantly different from zero or statistically significant at 1 per cent. The surplus of schooling coefficient is positive, which implies that overeducated workers still have economic value and receive positive returns to each additional year of overeducation. But, the coefficient of the surplus of schooling was slightly lower than the coefficient of years of required schooling. Furthermore, the coefficient of deficit of schooling is negative, meaning that undereducated workers receive lower returns compared to the returns received by matched workers in the same occupation. Also, the absolute value of the coefficient of deficit schooling is lower than that of required schooling. This finding is consistent with Duncan and Hoffman's (1981) and confirms that the highest return is received by workers with required schooling level relative to the other worker's categories. A possible explanation from the demand side is that rapid technological changes affect more companies in Indonesia to prefer workers with higher educational qualifications (Allen, 2016). From the supply side, undereducated workers' lack of education may restrict their productivities, given the complexity level of their jobs (Quintini, 2011).

Furthermore, the R squared is around 16 per cent. Also, the coefficients of the dummy year of 2007 and 2014 are positive; both of which are higher relative to the dummy year of 2000. This indicates that wage return increases wave after wave (7-year period) along with the improvement of economic growth in Indonesia from 2000 to 2014 period. As Yin (2016) asserts, coefficients of year dummies can be treated as indicators of the deepening of economic reforms and market improvements. As the dummy of time effect influences the estimation, the panel data model will accommodate these dummies.

Moreover, the present study also applies clustered standard errors in the OLS models; clustered by year and by individuals. The result is shown in Appendix XVII. The coefficient estimate results are similar with the pooled OLS results; the difference is in the values of the standard error. Thus, using the clustered standard errors would not influence the coefficient estimates but may affect the significance levels. It is worth noting that the standard errors in the pooled OLS are clustered by individuals. As a result, the results of the pooled OLS with dummy year and clustered standard errors by individuals are exactly the same. Although the pooled OLS estimation is consistent with previous empirical findings, it has been criticised due to the unobserved heterogeneity problem. As such, FE and RE models can be used to deal with the problem, as explained in the methods used here (Section 5.3.2).

Table 5.10: The ORU Model for All Individuals

Variable	Pooled OLS with dummies year			RE			FE		
	Coef	SD	P>t	Coef	SD	P>t	Coef	SD	P>t
Years of required schooling	0.106	0.005	***	0.101	0.006	***	0.047	0.022	*
Years of surplus schooling (OE)	0.100	0.006	***	0.101	0.006	***	0.060	0.022	**
Years of deficit schooling (UE)	-0.087	0.007	***	-0.088	0.007	***	-0.058	0.024	*
Experience	0.034	0.005	***	0.032	0.005	***	0.019	0.023	
Experience squared	-0.001	0.000	***	-0.001	0.000	***	0.000	0.000	*
Sex (1=female)	-0.310	0.025	***	-0.326	0.028	***		(omitted)	
Ethnicity (1=Javanese)	0.043	0.031		0.057	0.036			(omitted)	
Married and cohabitate	0.171	0.034	***	0.173	0.035	***	0.234	0.070	***
Other (Separated, divorced and widowed)	0.101	0.066		0.117	0.068		0.255	0.130	*
Status: full-time (30 hours a week or more)	-0.355	0.033	***	-0.366	0.033	***	-0.444	0.058	***
Tenure	0.073	0.005	***	0.069	0.005	***	0.036	0.008	***
Tenure squared	-0.002	0.000	***	-0.002	0.000	***	-0.001	0.000	***
Sector: private	-0.327	0.038	***	-0.309	0.040	***	-0.213	0.087	*
Industry2: mining and quarrying	0.490	0.110	***	0.393	0.112	***	0.094	0.206	
Industry3: manufacturing	0.153	0.047	**	0.181	0.047	***	0.149	0.091	
Industry4: electricity, gas and water	-0.171	0.135		-0.169	0.137		-0.287	0.241	
Industry5: construction	0.166	0.060	**	0.160	0.061	**	0.073	0.119	
Industry6: wholesale, retail, restaurants and hotels	0.008	0.051		0.036	0.052		0.036	0.101	
Industry7: transportation, storage, and communications	0.008	0.070		0.036	0.071		0.022	0.130	
Industry8: Finance, insurance, real estate and business services	0.391	0.074	***	0.383	0.074	***	0.186	0.142	
Industry9: Social services	0.057	0.046		0.087	0.047		0.033	0.091	
Firm size2: 20-99 people	0.219	0.028	***	0.204	0.029	***	0.072	0.050	
Firm size3: >= 100 people	0.422	0.035	***	0.386	0.035	***	0.137	0.063	*
Urban	0.132	0.027	***	0.146	0.028	***	0.091	0.067	
Capital region	0.260	0.038	***	0.255	0.043	***	0.287	0.197	
2007	0.085	0.030	**	0.095	0.028	***	0.2643	0.159	
2014	0.292	0.030	***	0.322	0.029	***	0.6471	0.3	*
Rho				0.438			0.630		
Constants	6.619	0.087	***	6.645	0.090	***	7.331	0.383	***
Number of observations	21,174			21174			21174		
Number of individuals				15440			15440		
The R-squared statistic	0.16								
Wald chi2(3)				3387.7		***			
F Test							29.84		***
Post Estimation:									
Breusch and Pagan Lagrangian multiplier test for random effects Var(u) = 0				31.270		***			
Hausman Test									
Ho: difference in coefficients not systematic							102.4		***

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

RE result shows that the wage return to an additional year of required and surplus schooling were the same (10.1 per cent). Meanwhile, the wage return to an additional year of deficit schooling was slightly different from the pooled OLS model. Yet, this result still has the same pattern as the previous models: workers with years of surplus schooling receive premium wages and workers with years of deficit schooling received penalty wages. The Rho (interclass correlation) was 43.8 per cent, indicating that 43.8 per cent of the variance is due to differences across cross-sections. The Wald test result was 3387.68 with a p-value of less than 0.01, this indicates that the model is adequate and the coefficients in the model are different from zero. Also, the Breusch-Pagan Lagrange multiplier (LM) result shows that chi2 was 31.3 with p-value = 0.00, meaning that the null hypothesis cannot be rejected, and RE is appropriate or more favourable than the pooled OLS model in this case.

Based on the FE result, the wage returns to additional year of surplus schooling were slightly higher than years of required schooling: 6 per cent and 4.7 per cent, respectively. Also, the coefficient of deficit of schooling was still negative at -5.8 per cent, which is higher than years of required schooling in absolute value. This is in contrast to Korpi and Tahlin (2007) and Dockery and Miller (2012) who find lower coefficients when applying FE, but it does not change the pattern of the coefficients. Some studies assert that the changes in the coefficient of schooling variables indicate that unobservable heterogeneity could affect the analysis substantially (Lindley and McIntosh, 2009 and Yin, 2013). However, explaining unobservable heterogeneity is not easy. Another possible reason is a collinearity with the fixed effect, for instance, several variables are constant over time for any given individuals, such as education (Wooldridge, 2009), so education may be collinear with the individual's level of fixed effects, as well as UE/OU/REQ status remains constant, only few individuals change the status. Similarly, year dummies are indeed the time fixed effect. Alternatively, the significant change in those coefficients could occur since the model for Indonesia's case is sensitive to the change in methods.

Having said that, Hausman test result shows that the chi2 was 102.4, with p-value = 0.00, which indicates that the FE model is more favourable than the RE. Yet, Bell and Jones (2015) argue that FE in this context is possibly not appropriate because the data used are longitudinal, *i.e.* the data could have a hierarchical structure; and the FE result is substantially different from the RE and the pooled OLS models. The education coefficients in FE estimates in this chapter are largely statistically insignificant, due to lack of variation over time in the education measure (this is time-invariant for most people). Thus, RE would be more appropriate and sensible, despite its more restrictive assumption that individual effect is uncorrelated with the exogenous variables. Bell and Jones also argue that the RE model is not simply technical solutions to endogeneity, but it is also the substantive importance of context/heterogeneity. In additions, Nielsen (2007) argues that only few individuals that change status for category of worker with required years of schooling to being surplus of years of schooling. Similarly, Wooldridge (2009) asserted some variables do not change over time such as education, as well as overeducated-required-undereducated statuses. Using fixed effect or first differencing, education variables cannot be included in the equation. Additionally, to test whether the return to education was constant over time, the interaction of education variable with year dummies can be used. Thus, based on those considerations, the present study prefers RE model for this case.

Based on the analysis, all models indicate that workers with years of surplus schooling receive premium wages. There are some explanations for this: Wye and Ismail (2018) assert that premium wage could be paid to overeducated workers because they are healthier, have stronger work and career aspiration, are more acceptable to on-the-job training, and have longer job tenure than perfectly matched workers in the same occupation (Büchel, 2002). These explain the persistent motive of employers to hire overqualified/overeducated workers with wage rewards without causing much wage inequality in the labour market (Rodriguez, 2011). Also, education in Indonesia might be used as a signal for ability rather than as a source of skill supply; and there is an increase in the demand for workers with higher levels of education in the country (Allen, 2016). There are also empirical data showing that there is a positive relationship between the education required and wages, and this will be elaborated in the sectoral analysis.

Furthermore, the present study performs the equality test for two coefficients as explained in the post-estimation test result: all models indicate that there was no statistical difference between coefficients of surplus and required years of education, including the FE model. RE model could provide clearer explanation that coefficients of surplus and required years of education are not statistical difference; since in terms of value, both coefficients have exactly the same value of 0.101, which implies workers with years of surplus may not receive lower wages than workers with required schooling. Meanwhile, the difference between coefficient of years of required schooling and deficit schooling was significant.

Also, penalty wages received by workers with years of deficit schooling is significantly different (Table 5.11).

Table 5.11: Testing the Equality of Two Coefficients

	Coefficient test		
	F-test (p-value)		
	Pooled OLS + dummies year	**RE**	**FE**
H0: years of required = years of OE	1.11 (0.29)	0.01 (0.93)	2.13 (0.14)
H0: years of required = years of UE	384.10 (0.00)	317.55 (0.00)	5.68 (0.02)

Source: The author's calculation.
Notes: p value is in brackets.

In short, all models have the same implications; undereducated workers receive negative or lower wages whereas overeducated workers still receive positive wage returns. This is also in line with Duncan and Hoffman's (1981) finding. The present study could suggest that overeducated workers still obtain higher or at least the same return as workers with required schooling. Thus, this promotes the education sector to expand rapidly. Though the RE result shows that the returns for workers with required and surplus schooling are the same, pursuing higher education requires more investment in education. Thus, from the workers' perspective, the highest return is still obtained by workers with required schooling.

5.4.2 Control Variables

The present study uses similar control variables as in Chapter 3, including some adjustments to the control variables such as: ethnicity (majority/non-majority), firm size (only 3 categories in this chapter), and capital region (adjusted from province dummies).

In terms of control variables, it seems that the result of the RE model (see Table 5.10) is similar to Chapter 3 (see Table 3.20). Firstly, for personal characteristic variables, sex is negative and significant at 1 per cent, implying that males receive higher hourly wages compared to females. This is in line with Becker's employer taste model of discrimination (Becker, 1971) and the statistical discrimination theory. Meanwhile, married and cohabitate status has a positive and significant effect on wages, as one of the wage elements of white-collar workers is family allowances (includes spouse and children according to the law number 8/1974) for married workers which inevitably increases the overall wage. In contrast, ethnicity has insignificant effects on wage. All these findings are similar to the arguments explained in Section 3.5.6.

Turning to work-related and firm size variables, experience and tenure variables have a positive impact on wage in any specifications, and the effects of the square of experience and tenure are negative, which is consistent with the hypotheses in the labour market (as explained in Part 3.2.2). Furthermore, for the full-time dummy variable, the result shows that all coefficients are negative and significant, which is still similar to Chapter 3 findings. The next control variable is sector; the private sector yields a negative and significant coefficient in any specifications, which also implies that an average government worker earns more, not less, than his/her private sector counterpart (World Bank, 2000). For the dummy of industries, only a few industries have a significant effect on wage, such as: mining and quarrying; manufacturing; construction; and finance, insurance, real estate and business services. Adding more data in the analysis strengthens the argument that Indonesia has a comparative advantage in labour-intensive manufacturing and that the service sector still dominates the economy (as explained in Chapter 2). The last variable in this category is firm size. The bigger the firms are, the higher the wages that the workers can earn, as is in line with Dhanani and Islam (2004).

With regards to residence, both urban/rural and capital/non-capital province dummies show positive and significant results. Thus, working in urban areas and/or in the capital province (DKI Jakarta) affects higher hourly wages, as is in line with Comola and de Mello (2011).

The additional dummy variables are year dummies. In both 2007 and 2014, these dummies are positive and significant relative to the dummy in 2000, which could imply that there is a deepening of economic reform and economic growth, as Yin (2016) asserts.

5.4.3 Wage Effects by Gender

In general, coefficients of years of required schooling, years of surplus schooling (OE), and years of deficit schooling (UE) of females are slightly higher than the coefficients of males (Table 5.12). This suggests that females have a higher return to required schooling, surplus schooling, as well as deficit schooling compared to males. It is worth noting that the interpretation of the coefficient estimates of the control variables are *ceteris paribus*.

Higher return for females also occurs in the Mincer wage equation (Chapter 3). Similar to the argument in Chapter 3 finding, a possible explanation for the higher return of females is the low probability or short duration of employment which requires a higher apparent rate of return compared to males (Abbas and Foreman-Peck, 2008). The empirical data on the average working week in Indonesia (Table 2.4) also support this argument; females tend to have shorter average working hours than males.

Based on the pooled OLS model for males, the wage return to additional years of required schooling was 10.2 per cent; the wage returns to one year of surplus schooling was 9.2 per cent; and the wage return to one year of deficit schooling was -8.2 per cent. Meanwhile, the RE model result for males indicates that the wage return to additional years of required schooling was 9.7 per cent; the wage return to one year of surplus schooling was 9.2 per cent; and the wage return to one year of deficit schooling was -8.1 per cent. The pattern of coefficients between the pooled OLS and RE was relatively similar. These results are consistent with Duncan and Hoffman's finding in 1981.

For females, the pooled OLS result shows that the wage returns to additional years of required schooling was 10.3 per cent; the wage returns to one year of surplus schooling was 10.7 per cent; and the wage returns to one year of deficit schooling was -9 per cent. Meanwhile, based on the RE model, the wage returns to additional years of required schooling, the wage returns to one year of surplus schooling, and the wage returns to one year of deficit schooling were 9.9 per cent, 10.7 per cent and -9.2 per cent, respectively. These results are slightly different from Duncan and Hoffman's (1981) finding, since the coefficients of years of surplus schooling are slightly higher than the coefficients of years of required schooling.

However, the coefficient's test result (Appendix XVIII) indicates the difference between coefficients of years of required schooling and surplus schooling was insignificant for both males and females in the pooled OLS and RE models. Meanwhile, the difference between coefficients of years of required schooling and deficit schooling was significant. The implication is similar to all sample results: workers with years of surplus schooling receive premium wages, which could be similar to premium wages received by workers with required schooling. Also, the penalty wages received by workers with years of deficit schooling was significantly different.

Using the FE approach, there are also some substantial changes in both the coefficient value as well as the patterns and significance of the variables: most mismatch coefficients were insignificant, except for years of surplus schooling of males (positive and significant at 10 per cent). Similar to all individuals, this could indicate that the model is very sensitive to the methods used. In particular, the female's FE result indicates that the coefficient of years of deficit schooling is far higher in absolute value than the coefficient of years of required schooling, though remains insignificant.

Females in Indonesia tend to take any jobs they can find, despite the low wages offered. As Alisjahbana and Manning (2006) find, better-off women are more likely to be unemployed and poorer women are more likely to be underemployed (working but wanting to work more). This shows that better-off women can afford to stay unemployed for longer periods while poorer women will take whatever work they can find, often in the agricultural and/or informal sectors. Moreover, poorer married women are more likely to participate in the labour market than married women in non-poor households. Thus, an unobserved heterogeneity factor for females is family background, and it may substantially affect the wage and mismatch relationship.

In short, the patterns of those coefficients of the pooled OLS and the RE model are relatively similar with all individuals estimation results. The implications are also the same for both males and females; overeducated workers receive wage premium and their over-qualification still have some economic values, or at least the wage could be similar to the premium wages received by workers with required schooling. In contrast, undereducated workers receive a wage penalty (lower wages) compared to the matched workers. Thus, increasing education attainment is necessary for both male and female workers. Comparing those three models, the RE model is preferred since it deals with heterogeneity and only few individuals that change status for those worker categories. similar reason with the main model. In additions, the present study conducted additional estimation with the interaction between sex and OE/REQ/UE/years of schooling. The interaction variables are significant affect the wages, with the exception Model 4, the interaction between sex and REQ is insignificant. Thus, the effect of years of deficit and surplus schooling on wages would be different at different gender. The estimations are provided in Appendix XVI Table XVI.4.

Table 5.12: The ORU Model by Gender

Variable	Male									Female								
	Pooled OLS with dummies year			RE			FE			Pooled OLS with dummies year			RE			FE		
	Coef	SD	P>t	Coef	SD	P>t	Coef	SD	P>t	Coef	SD	P>t	Coef	SD	P>t	Coef	SD	P>t
Years of required schooling	0.102	0.007	***	0.097	0.007	***	0.049	0.027		0.103	0.008	***	0.099	0.009	***	0.041	0.038	
Years of surplus schooling (OE)	0.092	0.007	***	0.092	0.008	***	0.064	0.027	*	0.107	0.010	***	0.107	0.011	***	0.049	0.039	
Years of deficit schooling (UE)	-0.082	0.009	***	-0.081	0.009	***	-0.047	0.029		-0.090	0.011	***	-0.092	0.012	***	-0.073	0.041	
Experience	0.039	0.006	***	0.040	0.006	***	0.044	0.028		0.026	0.008	***	0.021	0.008	**	-0.035	0.040	
Experience squared	-0.001	0.000	***	-0.001	0.000	***	-0.001	0.000	**	-0.001	0.000	**	0.000	0.000	*	0.000	0.000	
Ethnicity (1=Javanese)	0.057	0.040		0.064	0.046			(omitted)		0.023	0.051		0.053	0.059			(omitted)	
Married and cohabitate	0.240	0.044	***	0.234	0.045	***	0.207	0.088	*	0.098	0.056		0.110	0.057		0.353	0.117	**
Other (Separated, divorced and widowed)	0.208	0.112		0.241	0.112	*	0.295	0.189		0.053	0.087		0.064	0.090		0.276	0.183	
Status: full-time (30 hours a week or more)	-0.493	0.045	***	-0.494	0.045	***	-0.530	0.079	***	-0.239	0.048	***	-0.263	0.048	***	-0.349	0.084	***
Tenure	0.056	0.006	***	0.052	0.006	***	0.023	0.009	*	0.099	0.008	***	0.095	0.008	***	0.061	0.013	***
Tenure squared	-0.001	0.000	***	-0.001	0.000	***	0.000	0.000		-0.002	0.000	***	-0.002	0.000	***	-0.002	0.000	***
Sector: private	-0.292	0.048	***	-0.269	0.051	***	-0.040	0.112		-0.372	0.062	***	-0.366	0.066	***	-0.483	0.137	***
Industry2: mining and quarrying	0.493	0.114	***	0.398	0.117	***	0.111	0.216		0.795	0.483		0.764	0.487		0.568	1.179	
Industry3: manufacturing	0.202	0.057	***	0.227	0.057	***	0.180	0.109		0.115	0.084		0.134	0.084		0.039	0.166	
Industry4: electricity, gas and water	-0.211	0.142		-0.195	0.145		-0.271	0.266		0.592	0.465		0.303	0.439		-0.532	0.618	
Industry5: construction	0.137	0.065	*	0.129	0.066		0.057	0.132		0.335	0.199		0.353	0.198		0.146	0.334	
Industry6: wholesale, retail, restaurants and hotels	0.071	0.063		0.090	0.064		0.031	0.123		-0.023	0.090		-0.002	0.090		0.025	0.179	
Industry7: transportation, storage, and communications	-0.020	0.075		0.018	0.076		0.029	0.143		0.340	0.234		0.252	0.231		-0.216	0.424	
Industry8: Finance, insurance, real estate and business services	0.393	0.087	***	0.378	0.088	***	0.091	0.165		0.456	0.137	***	0.465	0.140	***	0.539	0.284	
Industry9: Social services	0.064	0.056		0.097	0.057		0.011	0.109		0.106	0.082		0.116	0.083		0.087	0.169	
Firm size2: 20-99 people	0.165	0.035	***	0.160	0.036	***	0.076	0.062		0.319	0.047	***	0.286	0.048	***	0.081	0.087	
Firm size3: >= 100 people	0.364	0.044	***	0.317	0.044	***	0.062	0.077		0.547	0.057	***	0.531	0.058	***	0.321	0.112	**
Urban	0.131	0.033	***	0.131	0.035	***	-0.021	0.082		0.151	0.045	***	0.196	0.048	***	0.351	0.117	**
Capital region	0.180	0.049	***	0.170	0.054	**	0.101	0.221		0.395	0.063	***	0.392	0.070	***	1.144	0.445	*
2007	0.070	0.038		0.076	0.034	*	0.158	0.193		0.123	0.052	*	0.135	0.046	**	0.482	0.284	
2014	0.251	0.037	***	0.282	0.036	***	0.448	0.364		0.382	0.050	***	0.398	0.048	***	1.044	0.532	*

Constants	6.758	0.107	***	6.765	0.111	***	7.316	0.486	***	6.164	0.146	***	6.204	0.151	***	7.236	0.634	***
Rho				0.415			0.616						0.478			0.665		
Number of observations	13,312			13312			13312			7,862			7862			7862		
Number of individuals				9539			9539						5901			5901		
The R-squared statistic	0.14									0.19								
Wald chi2(3)				1796.90		***							1570.60		***			
F Test							17.83		***							14.06		***
Post Estimation:																		
Breusch and Pagan Lagrangian multiplier test for RE																		
Var(u) = 0				17.320		***							12.450		***			
Hausman Test																		
Ho: difference in coefficients not systematic							66.17		***							65.12		***

Source: The author’s calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

5.4.4 Wage Effects by Sector

Turning to wage effects by sector, the conclusion is still the same with the main model (Table 5.13). The interpretation of the coefficient estimates of the control variables are *ceteris paribus*. Meanwhile, the FE result shows that most of the coefficients are insignificant with the exception of years of surplus schooling of the private sector. This is possibly due to collinearity with the fixed effect, as explained in Section 5.4.1.

In terms of the coefficients in RE model, it seems that the coefficients of years of surplus schooling are slightly higher than the coefficients of years of required schooling. But the equality of two coefficient test results show that those are insignificant (Appendix XVIII).

Moreover, there are some empirical data in relation to wages and education in both sectors. In the public sector, entry ranks are mainly determined by education level, and increases in rank are largely driven by seniority, with a maximum rank depending on the entry level of the civil servant (World Bank, 2002). For instance, an undergraduate entry level is IIIA and master's degree is IIIB. In terms of wages, IIIB with 0-year experience receives a slightly higher wage (around GBP 7 per month) compared to IIIA with 0-year experience (Government Regulation 30/2015). Similarly, most private companies apply different wage rates based on educational qualifications. This indicates that there is a positive relationship between years of schooling and wage in general, for both the public and the private sectors.

Comparing the coefficients of the mismatch, most of the coefficients are higher in the public sector than in private sector. However, noncompeting groups model is suitable for Indonesia as one of developing countries, in which individuals belong to one labour market segment or another, and they cannot or will not switch from one to another (Fields, 2010). This finding is consistent with Chapter 3 finding that the return to education in the public sector is higher than in the private sector. The distribution of earnings across the public sector suggests that government wage-setting institutions are very different from those in the private economy. The lack of competition in the public sector also allows higher returns to education in the sector (Psacharopolous, 1979).

The public sector also has higher coefficient of years of surplus schooling. This is in line with Allen *et al*. (2013) who find that overeducated workers in the public sector are likely to be paid higher wages than one would expect based on their productivity, based on wage setting institution theory (as explained in Section 5.2.5). When applying FE, the result shows some substantial changes in both the coefficient value and patterns and significance of variables. Most mismatch coefficients were insignificant, except for years of surplus schooling of the private sector (positive and significant at 10 per cent). These findings may indicate that the ORU model in this study is relatively sensitive to the panel data method used.

Table 5.13: The ORU Model by Sector

Variable	Private									Public								
	Pooled OLS with dummies year			RE			FE			Pooled OLS with dummies year			RE			FE		
	Coef	SD	P>t	Coef	SD	P>t	Coef	SD	P>t	Coef	SD	P>t	Coef	SD	P>t	Coef	SD	P>t
Years of required schooling	0.091	0.006	***	0.090	0.006	***	0.049	0.027		0.145	0.012	***	0.127	0.013	***	-0.018	0.041	
Years of surplus schooling (OE)	0.091	0.007	***	0.091	0.007	***	0.054	0.027	*	0.141	0.013	***	0.132	0.013	***	0.004	0.041	
Years of deficit schooling (UE)	-0.075	0.008	***	-0.076	0.008	***	-0.052	0.028		-0.123	0.019	***	-0.127	0.019	***	-0.010	0.048	
Experience	0.025	0.005	***	0.025	0.005	***	0.038	0.028		0.047	0.011	***	0.045	0.011	***	0.001	0.043	
Experience squared	-0.001	0.000	***	-0.001	0.000	***	-0.001	0.000	*	-0.001	0.000	*	0.000	0.000		0.000	0.001	
Sex (1=female)	-0.355	0.028	***	-0.360	0.031	***	(omitted)			-0.115	0.051	*	-0.133	0.061	*	(omitted)		
Ethnicity (1=Javanese)	0.045	0.035		0.060	0.039		(omitted)			0.066	0.071		0.092	0.086		(omitted)		
Married and cohabitate	0.192	0.038	***	0.192	0.038	***	0.232	0.081	**	0.267	0.087	**	0.230	0.090	*	0.188	0.176	
Other (Separated, divorced and widowed)	0.137	0.073		0.147	0.075	*	0.286	0.152		0.130	0.156		0.141	0.160		0.156	0.271	
Status: full-time (30 hours a week or more)	-0.382	0.038	***	-0.403	0.038	***	-0.517	0.076	***	-0.305	0.060	***	-0.295	0.059	***	-0.429	0.091	***
Tenure	0.086	0.006	***	0.080	0.005	***	0.037	0.009	***	0.052	0.010	***	0.049	0.010	***	0.032	0.015	*
Tenure squared	-0.002	0.000	***	-0.002	0.000	***	-0.001	0.000	**	-0.001	0.000	**	-0.001	0.000	**	-0.001	0.000	*
Industry2: mining and quarrying	0.449	0.119	***	0.357	0.120	**	-0.044	0.239		0.619	0.298	*	0.533	0.305		0.087	0.527	
Industry3: manufacturing	0.168	0.050	***	0.191	0.050	***	0.189	0.101		-0.377	0.187	*	-0.297	0.186		-0.302	0.325	
Industry4: electricity, gas and water	-0.152	0.157		-0.147	0.157		-0.152	0.290		-0.288	0.259		-0.442	0.267		-1.328	0.477	**
Industry5: construction	0.208	0.063	***	0.207	0.064	**	0.131	0.130		-0.858	0.240	***	-0.824	0.243	***	-0.341	0.417	
Industry6: wholesale, retail, restaurants and hotels	0.032	0.054		0.050	0.055		0.074	0.112		-0.590	0.259	*	-0.427	0.261		0.048	0.511	
Industry7: transportation, storage, and communications	0.042	0.074		0.066	0.075		0.091	0.144		-0.351	0.214		-0.168	0.216		0.000	0.385	
Industry8: Finance, insurance, real estate and business services	0.422	0.080	***	0.410	0.080	***	0.305	0.160		0.154	0.206		0.219	0.211		-0.010	0.404	
Industry9: Social services	0.063	0.050		0.082	0.051		0.042	0.105		-0.088	0.125		0.005	0.125		0.138	0.205	
Firm size2: 20-99 people	0.256	0.033	***	0.231	0.033	***	0.060	0.061		0.099	0.052		0.124	0.055	*	0.134	0.092	
Firm size3: >= 100 people	0.471	0.039	***	0.433	0.039	***	0.199	0.075	**	0.346	0.075	***	0.315	0.076	***	0.006	0.132	
Urban	0.147	0.031	***	0.160	0.032	***	0.141	0.081		0.148	0.052	**	0.152	0.057	**	-0.034	0.126	
Capital region	0.287	0.041	***	0.278	0.045	***	0.324	0.218		-0.028	0.110		-0.004	0.126		-0.042	0.548	
2007	0.088	0.035	*	0.085	0.032	**	0.090	0.197		0.058	0.062		0.084	0.055		0.582	0.281	*

2014	0.327	0.034	***	0.336	0.033	***	0.335	0.375		0.130	0.064	*	0.205	0.060	***	1.206	0.516	*
Constants	6.479	0.079	***	6.507	0.081	***	6.959	0.426	***	5.962	0.200	***	6.069	0.207	***	8.630	0.910	***
Rho				0.423			0.625						0.496			0.702		
Number of observations	17,512			17,512			17,512			3,662			3662			3662		
Number of individuals				13417			13417						2429			2429		
The R-squared statistic	0.12									0.18								
Wald chi2(3)				2171.14		***							723.86		***			
F Test							18.24		***							12.86		***
Post Estimation:																		
Breusch and Pagan Lagrangian multiplier test for RE																		
Var(u) = 0				16.200		***							14.970		***			
Hausman Test																		
Ho: difference in coefficients not systematic							73.84		***							82.91		***

Source: The author’s calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

5.4.5 Robustness Test

The present study uses three approaches as the robustness test, *i.e.* replacing mode in the main model by mean, using balanced panel data, and applying an alternative model from Verdugo and Verdugo (1989). Those tests confirm that overeducated workers receive wage premium for each year of surplus of schooling, while undereducated workers receive wage penalty for each year of deficit schooling. The coefficients of deficit schooling are negative but insignificant in affecting the wages. Meanwhile, using Verdugo and Verdugo's model, the present study finds a different conclusion compared to Verdugo and Verdugo (1989) results. Previous empirical studies find that overeducated workers will receive penalty wage and undereducated workers will receive premium wage, such as: Iriondo and Perez-Amaral (2013) and Alisjahbana *et al.* (2017). The present study finds that overeducated workers may receive wages higher than the matched workers. Also, undereducation incidences may not have any effect on wage. However, this is not the main focus of the analysis here and this issue can be explored in future research.

Realised Method by Mean

When comparing the RE models between mode (Table 5.10) and mean (Table 5.14), the results still show a similar pattern in coefficients of the main variables and thus suggest the same implications; the surplus of schooling coefficient is positive and significant and the deficit of schooling coefficient is negative and significant on wage. This implies that one extra year of surplus schooling (overeducation) still gains premium wages, whereas one extra year of deficit schooling (undereducation) will receive penalty wages. When comparing premium wages between required schooling and surplus schooling, the result shows that workers with required schooling receive higher wages than workers with surplus schooling. Similarly, the return to workers with required schooling is still higher than the absolute value of the return to undereducated workers. This agrees with Duncan and Hoffman's (1981) finding. The complete results and sample distribution are provided in Appendix XIX.

The RE model also shows a similar pattern with the pooled OLS model. There is a slight difference between the mode (main model) and the mean results: the coefficients of years of required schooling and surplus schooling are the same in the model. The t-test also indicates an insignificant difference between those coefficients. In the mean model, those coefficient values are different: the coefficient of required schooling is 12.1 per cent while the coefficient of surplus schooling is 9 per cent.

Controlling unobservable heterogeneity, the FE model shows some substantial changes. All mismatch coefficients are insignificant and the absolute value of years of deficit schooling is slightly higher than the coefficient of years of required schooling. Using mean also confirms that the model is sensitive to the method of panel data used, as explained in Section 5.3.2.

Table 5.14: Estimation Result Based on Mean

Variable	Pooled OLS with dummies year			RE			FE		
	Coef	SD	P>t	Coef	SD	P>t	Coef	SD	P>t
Years of required schooling	0.133	0.007	***	0.121	0.008	***	0.012	0.019	
Years of surplus schooling (OE)	0.096	0.007	***	0.090	0.007	***	-0.002	0.018	
Years of deficit schooling (UE)	-0.076	0.008	***	-0.073	0.009	***	-0.019	0.019	
	…	…		…	…		…	…	
Constants	6.372	0.100	***	6.484	0.104	***	8.035	0.309	***
Rho				0.438			0.652		
Number of observations	21,174			21,174			21,174		
Number of individuals				15,440			15,440		
The R-squared statistic	0.1557								
Wald chi2(3)				3298.850		***			
F Test							29.49		***
Post Estimation:									
Breusch and Pagan Lagrangian multiplier test for RE									
Var(u) = 0				29.840		***			
Hausman Test									
Ho: difference in coefficients not systematic							140.9		***

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Balanced Panel Data

The other approach of robustness test is using balanced panel data. The main aim of this approach is to compare the result of balance data with non-random attrition and the unbalance panel data with random attrition, whether the results still indicates similar pattern.

This approach is shrinking the sample to only individuals who appear in all three waves of the survey, which means the later waves are skewed towards older age groups. As indicated in Table 5.15, there pattern of the mean of year of surplus schooling in balance and unbalance panel data is not significantly different based on t-test result (Table 5.15). On the other side, the trend of years of deficit of balanced panel data is significantly higher than the trend in unbalance panel data, which may indicate older generation has a significant contribution on the increase of undereducation (workers with deficit years of schooling).

Table 5.15: Comparison of the Mean of Years of Surplus Schooling and Years of Deficit Schooling, Balanced and Unbalanced Panel Data

	2000	2007	2014
Years of Surplus Schooling			
Balanced panel	29.43	20.92	20.3
Unbalanced panel	28.5	24.9	23.4
Two tails t-test (P-Value)	0.3078		
Years of Deficit Schooling			
Balanced panel	14.27	25.35	28.01
Unbalanced panel	13.6	20.6	22
Two tails t-test (P-Value)	0.1418		

Source: The author's calculation.
Notes: T-test hypothesis is mean of balanced panel data is equal to mean of unbalanced panel data.

Although the attritions between this robustness test and the main model are different; and there is skewed towards older age groups in the robustness test, this approach also does not change the key findings of the present study. Overeducated workers and workers with required years of schooling still receive premium wages and undereducated workers still receive penalty wages. The pattern of coefficients is similar to the main model (all individuals). The pooled OLS and RE results are more consistent to Duncan and Hoffman's (1981) finding that the wage return to one-year surplus schooling is lower than the wage returns to one-year required schooling. Meanwhile, the wage return to one-year deficit schooling is less than the wage returns to required schooling in absolute value. Controlling unobservable heterogeneity, the FE result significantly changes: the wage return to years of surplus schooling is higher than the wage returns to years of required schooling; and the absolute value of years of deficit schooling is higher than the wage returns of required schooling. Again, this finding confirms that the model is sensitive to the method of panel data used. This finding may also imply that unobservable heterogeneity (such as ability) affects the model. Thus, there is no significant difference in pattern of variables between unbalanced and balanced panel data. The complete result, summary statistics and sample distribution are provided in Appendix XX.

Table 5.16: Estimation Result, Balanced Panel

Variable	Pooled OLS with dummies year			RE			FE		
	Coef	SD	P>t	Coef	SD	P>t	Coef	SD	P>t
Years of required schooling	0.119	0.010	***	0.119	0.010	***	0.080	0.032	*
Years of surplus schooling (OE)	0.105	0.011	***	0.107	0.012	***	0.084	0.032	**
Years of deficit schooling (UE)	-0.113	0.012	***	-0.116	0.013	***	-0.092	0.034	**
…	…	…		…	…		…	…	
Constants	7.172	0.187	***	7.148	0.194	***	7.673	0.705	***
Rho				0.084			0.352		
Number of observations	3,384			3,384			3,384		
Number of individuals				1,128			1,128		
The R-squared statistic	0.2681								
Wald chi2(3)				1122.92		***			
F Test							15.01		***
Post Estimation:									
Breusch and Pagan Lagrangian multiplier test for RE									
Var(u) = 0				21.690		***			
Hausman Test									
Ho: difference in coefficients not systematic							71.3		***

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Alternative model: Verdugo and Verdugo (1989)

In Chapter 4, the present research finds that there were the increase in undereducation and the decrease in overeducation during 2000-2014 period. Table 5.16 also confirms these findings: there was an increase in sample proportion for dummy undereducation equal to one, from 13 per cent in 2000 to 20.9 per cent in 2007 and to 22.7 per cent in 2014. There was also a decrease in sample proportion of dummy overeducation equal to one, from 28.5 per cent in 2000 to 24.9 per cent in 2007 and to 23.3 per cent in 2014.

Table 5.17: Sample Distribution for Overeducation and Undereducation Dummies

		Years		
		2000	2007	2014
Dummy of undereducation	**0**	86.4	79.1	77.3
	1	13.6	20.9	22.7
	Total	100.0	100.0	100.0
Dummy of overeducation	**0**	71.5	75.1	76.7
	1	28.5	24.9	23.3
	Total	100.0	100.0	100.0

Source: The author's calculation.

By applying the Verdugo and Verdugo model (as explained in the literature review); the present study finds that the coefficients of dummy of overeducation are positive and significant at 1 per cent for the pooled OLS model; positive and significant at 5 per cent for the RE model, and insignificant in the FE model. Positive and significant coefficient implies that overeducated workers receive premium wages relative to workers with actual years of schooling. In terms of the dummy of overeducation, the coefficient is 11.8 per cent for the pooled OLS model. When applying the RE model, the coefficient decreases slightly to 10.9 per cent. This implies that overeducated workers may receive 11.8 per cent higher wage compared to matched workers. Meanwhile, the coefficient of undereducation dummy is negative but insignificant in all specifications, which implies that undereducation incidences may not have any effects on wage. The complete estimation results of this model are provided in Appendix XXI.

This finding is inconsistent with Verdugo and Verdugo's (1989) findings and previous empirical studies which use this model: the coefficient of overeducation is negative and significant, whereas the coefficient of undereducation is positive and significant. This implies that overeducated workers receive wage penalty, relative to matched workers. Finally, this finding also confirms that the model is relatively sensitive to the panel data method used since the result of the RE and FE model are significantly different.

Table 5.18: Estimation Result Based on Verdugo and Verdugo's Model

Variable	Pooled OLS with dummies year			RE			FE		
	Coef	SD	P>t	Coef	SD	P>t	Coef	SD	P>t
Years of actual schooling	0.112	0.005	***	0.109	0.006	***	0.059	0.033	
Dummy of Overeducation	0.118	0.034	***	0.094	0.033	**	-0.041	0.055	
Dummy of Undereducation	-0.034	0.030		-0.016	0.030		0.026	0.051	
…	…	…		…	…		…	…	
Constants	6.548	0.088	***	6.564	0.092	***	7.177	0.556	***
Rho				0.437			0.628		
Number of observations	21,174			21,174			21,174		
Number of individuals				15,440			15,440		
The R-squared statistic	0.159								
Wald chi2(3)				3406.81		***			
F Test							29.64		***
Post Estimation:									
Breusch and Pagan Lagrangian multiplier test for RE									
Var(u) = 0				30.860		***			
Hausman Test									
Ho: difference in coefficients not systematic							98.73		***

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

5.5 Conclusion

Throughout the literature on education mismatch and wage in Indonesia, most studies have focused on analysing the relationship between wage and education mismatch by using cross-sectional data. The present study contributes to the discussion on education mismatch by using the ORU model which is relatively more systematic and more comprehensive than previous empirical studies (Table 5.1), since the model accommodates many variables that are grouped by personal characteristics, job related, firm size and region categories. In terms of education level, the present study covers from 6 years of schooling (primary school) to the highest level, since the challenge of the Indonesian labour market is the domination of workers with lower level of education in the labour market, and the increase of mismatch which is driven by the increase of undereducation (Chapter 4 finding). Thus, undereducation should be a deeper concern here than in developed countries since there is a risk of weaker productivity growth and a slower structural transition to higher value-added activities with undereducated workforce (ILO, 2015). This study also attempts to address the unobserved heterogeneity issue. Gender and sectoral differences are also considered to provide a more comprehensive and detailed analysis of the labour market.

This chapter has focused on answering three research questions: (5.1) does education mismatch (undereducation and/or overeducation) contribute to determine wage in Indonesia? (5.2) considering the unobserved heterogeneity, does education mismatch still contribute to determine wage, and (5.3) do returns associated with education mismatch differ by gender and by sector? Three waves of IFLS data (2000, 2007 and 2014) are used in the analysis. The present study employs the ORU model adapted from Duncan and Hoffman (1981) with panel analysis. The model offers a better interpretation of mismatch and wage relationship, in terms of each additional deficit or surplus of education. Meanwhile, panel data are used to deal with unobserved heterogeneity problem.

Answering the first research question, the results reported here indicate that both overeducation and undereducation incidences affect the labour wages in Indonesia. Based on the pooled OLS model, the present study finds that the wage returns to additional years of required schooling was around 10.6 per cent. The wage returns to one year of surplus schooling was around 10 per cent and the wage return to one year of deficit schooling was -8.7 per cent. This finding is consistent with Duncan and Hoffman's (1981) finding that: the wage return to one year of surplus schooling is positive and significant but lower than the wage return to additional years of required schooling; and the wage return to one year of deficit schooling is negative and significant with the absolute value lower than the wage return to an additional year of required schooling, However, the test of equality of two coefficients shows that the difference between the coefficients of years of required schooling and surplus schooling is insignificant. Also, the penalty wages received by workers with years of deficit schooling is significantly different from the wage return to additional years of required schooling.

The panel data model is used to deal with unobserved heterogeneity; this part also answers the second research question. When applying the RE, the result becomes similar to the pooled OLS findings. The results suggest that workers with years of surplus schooling receive premium wages and workers with years of deficit schooling receive penalty wages, which is consistent with Duncan and Hoffman's (1981) findings. Meanwhile, the

FE method's result indicates a substantial change in the coefficients value, pattern and significance level.

Furthermore, the post-estimation test indicates that FE is favourable. On the other hand, Bell and Jones (2015) argue that RE is more favourable in the context of longitudinal data and when hierarchies occur in the data. Considering the sensitive estimation results when using panel data methods, a cautious interpretation is needed in this case. Significantly different results of the RE and FE may indicate that the ORU model in this study is relatively sensitive to the panel data methods used. Another possible explanation is a collinearity with the fixed effect; both the individual's level fixed effects and the time fixed effects. Alternatively, the unobserved heterogeneity may affect the wages. Yet, obtaining the proof is not straightforward. Hartog (2000) asserts that the omission of unobserved heterogeneity may lead to an underestimation/overestimation of the rate of return to education mismatch. Meanwhile, Wen and Maani's (2017) argue that applying panel data techniques to the study of overeducation may lead to the controversial results on whether or not education and occupational mismatch with earning penalty effects are verified once individual heterogeneity is taken into account.

Answering the third research question, the coefficients of years of required schooling, years of surplus schooling (OE), and years of deficit schooling (UE) of females are slightly higher than those coefficients of males, which implies that females have a higher return to required schooling, surplus schooling, as well as deficit schooling than males do. This could be due to the short duration of females' employment (Abbas and Foreman-Peck, 2008). Similar to all individuals, the panel data method may result in a very sensitive result of estimation of both the RE and FE models. The implication of the gender analysis is similar to all sample results: both male and female workers with years of surplus schooling receive premium wages. The premium could be similar to premium wages received by workers with required schooling. Meanwhile, penalty wages received by workers with years of deficit schooling are significantly different.

By sector, the same key finding is also found within the general (all individuals) results. Workers in the public sector receive slightly higher wages than those in the private sector for workers with required and surplus years of schooling, as is in line with Allen *et al.* (2013). Also, overeducated workers in the public sector are likely to be paid higher wages than one would expect based on their productivity based on wage setting institution theory, and higher returns in the public sector support some signalling value of education

(Chevalier *et al.*, 2004). Applying the panel data method, the RE result is relatively similar to the pooled OLS result, meanwhile the FE result shows some substantial changes in both the coefficient value as well as patterns and significance of variables. Again, the models are relatively sensitive to the panel data method used.

This finding has several implications: education mismatch, which has negative effects on both the individual and the macro levels such as lower productivity. In this case, the workers acquire knowledge that is not subsequently transferred into skills that are needed for a certain job. This also highlights some inefficiency in the country's educational system and labour market. The policy makers can provide labour market information including job vacancies. Moreover, the government should focus on increasing education attainment, in particular senior high school and university levels to meet the growing demand for skills/education in the labour force. Equal access to senior high schools and universities is necessary in terms of physical infrastructure (higher education building construction) as well as funding mechanisms (such as scholarship and student loan). The other important aspect is that the proliferation of education in Indonesia should be accompanied by progress in the quality of schooling, such as: strengthening vocational education and skill training that can help prepare the youth for a smooth transition to employment; and promoting the integration of internship programs and other practical learning experience in the curriculums.

Owing to the limitations of the dataset, the present study is unable to elaborate the aspects of skill, subject of study and job characteristics. These data would allow the analysis to be more comprehensive for both the vertical and horizontal mismatches and to elaborate the job characteristics and skills required from the demand side. In the absence of other measures of skills competencies, the results from this analysis should not be treated as conclusive, but rather as indicative. The resulting skill gap is worrying as it will affect Indonesia's competitiveness in the world economy, which is becoming more and more knowledge-based and technologically intensive. Moreover, adding interaction variables such as gender and OE/REQ/UE variables will be more interesting and enrich the analysis of gender different. Finally, the assessment of the changes in the consequences of education mismatch in the labour market over time is an interesting area to be explored in future research.

Chapter 6 Conclusion and Implications

This chapter comprises three parts: (6.1) an overview of the key findings from each chapter; (6.2) the implications of the research and the policy recommendations related to return to education and education mismatch in Indonesia; and finally, (6.3) the research limitations and extensions as well as some points which may be worth further investigation in future research.

6.1 An Overview of Key Findings

This thesis focuses on exploring education and the labour market outcomes in Indonesia. It provides an overview of the education system and development as well as the background of the data used in Chapter 2. Meanwhile, the main discussion analyses three aspects, *i.e.*: the return to education (Chapter 3); the determinants of education mismatch (Chapter 4); and the relationship between education mismatch and wage (Chapter 5). This overview is drawn from the Indonesia Family Life Survey's (IFLS) data of 2000 and 2014 periods. Those waves are chosen to analyse the return to education and the education mismatch determinants in Chapter 3 and Chapter 4. Those waves represent the conditions before and after the education expansion period; hence the immediate and longer effects of those policies can be addressed. The IFLS of 2007 data are also added to increase the number of individuals and to narrow the gap between the periods of the analysis in Chapter 5 (education mismatch and wage). Compared to other data such as SAKERNAS (Labour Force Survey), IFLS data are more comprehensive since they can accommodate all variables that are needed in the present study, especially the data on the public and private workers.

Indonesia is the world's fourth most populous country in the world and one of the largest economies in Southeast Asia. As such, human capital is supposed to be one of the primary focuses in developing the country. However, the number of people with junior high school or lower qualifications in the country is around 64 per cent of the total population in 2017, which is still far higher than the number of those with senior high school or higher qualifications (Statistics Indonesia, 2019).

Education becomes one of the most important factors in career progression, as well as in increasing wage and productivity at both the individual and aggregate levels. The education sector then expands as many individuals become aware of the importance of education and decide to pursue higher education qualifications. In the US, such expansion is followed by the rise of overeducation incidences, since the supply of university graduates increases more rapidly than its demand (Freeman, 1976). This begs the question: does this also apply in developing countries?

In detail, Chapter 3 highlights several key points: the return to education varies and increases in line with education level, as predicted by the human capital theory (Schultz, 1961; Mincer, 1974; and Becker, 1994). Based on the Mincer wage equation, the return to junior high school in 2000 was 28.5 per cent, senior high school was 63.7 per cent and university was 119.3 per cent (relative to primary school and below). In 2014, the return to junior high school was 22.1 per cent and to senior high school and university were 48 per cent and 95.8 per cent, respectively. In terms of years of schooling, the result shows that one additional year of schooling increased the wage by 12.4 per cent in 2000 and by 9.9 per cent in 2014. In comparison, the return to junior high school in China in 2005 was 15.9 per cent; the return to senior high school was 46.6 per cent and the return to university was 95.8 per cent - relative to primary school (Gropello and Sakellariou, 2010). Indonesia's return to education in both periods was slightly higher, but the growth of the return was negative or declining. In contrast, China still had a positive growth of return to schooling for all education levels, as indicated in Table 3.5 (Gropello and Sakellariou, 2010).

The present study also finds a gender disparity in wage; the return to education for females is higher than for males. This trend also occurred in 139 other countries between the period of 1950 and 2014 (Psacharopoulos, 2018). Some possible explanations are the more limited supply of skilled female workers, the different technological requirements in female-dominated and male-dominated jobs (Ren and Miller, 2012), and shorter duration of females employment. A similar trend also occurs in other developing countries such as Pakistan (Abbas and Foreman-Peck, 2008).

In terms of sector, the return to education was generally higher in the public sector than in the private sector for both periods. Thus, the public sector in Indonesia continues to attract more job seekers. Some possible reasons for the premium in the public sector are: the increase of wage in public sector is higher and faster (The World Bank, 2000); there is a lack of competition in the public sector (Psacharopolous, 1979); the private sector is more efficient than the public sector (Rao, 2015); and wage rigidity may occur more in the public sector.

For trend over time, based on the equality of coefficients across two Mincer wage equations, the present study finds that the return to education tends to decrease for most education levels from 2000 to 2014. This decline could be attributed to the education expansion. This is consistent with the increase in the supply of educated workers (Section 3.2.4). This is also similar to Freeman's finding that there is a fall in the return to education in the US, which is attributed to the expansion of the country's education sector (Freeman, 1976). The same pattern also occurred in the UK between 1994 and 2006 as Walker and Zhu (2008) find. Furthermore, the decline in the return to higher education was because the higher education participation rate increased dramatically while the growth in relative labour demand suggests that the supply of graduates considerably outstripped the demand, which ought to imply a fall in the wage premium. Dumauli (2015) adds that one of the possible reasons for the decrease in the return to education in Indonesia is the low quality of the education system.

By gender, the highest decrease in the return to education is the return for males with junior high school qualifications (relative to primary school or below), one of possible explanations is more males having higher education attainment due to the 9-year (primary and junior high schools) compulsory education program. On the demand side, the economy transforms to service and manufacturing and increases the demand for female workers; for instance, garment companies hardly recruit men because the job needs sewing; moreover, an increase in retail sector causes supermarkets and mega-malls (which requires at least senior high school qualification) to replace the traditional markets. Thus, males with lower education attainment are more affected.

By sector, the private sector experienced a significant decline in the return to education between the year 2000 and 2014. Meanwhile, according to the test of coefficients of the two models, the result indicates that the changes are insignificant for the public sector. This is possibly due to the wage rigidity in the public sector as the sector has a wage mechanism that is not based on the market's mechanism.

Anticipating endogeneity problems and selection bias, the Instrumental Variable (IV) and Heckman model are used by incorporating conventional, policy and alternative instruments. The IV is used to gauge the role of the omitted variables (ability bias) in the OLS estimates of the return to schooling in the Indonesian labour market. The selection bias, in this case, is the bias that arose from non-random sampling for employment. The OLS with years of schooling is used to simplify the analysis. The OLS result shows that the coefficient of years of schooling in 2000 was 12.4 per cent; and decreased to 9.9 per cent in 2014. The IV results show that endogeneity problems may occur. In spite of the instruments' weaknesses, the IV result confirms that there is a decline in year schooling coefficient from 0.108 in 2000 to 0.066 (Model 1) and 0.088 (Model 2) in 2014. While the Heckman model confirms that selection bias could occur in the OLS model. But the conclusion remains the same: there is a decline in return to education, from 11.8 per cent in 2000 to 9.8 per cent in 2014.

In Chapter 4, another consequence of education expansion is overeducation, as has happened in the US. In developing countries such as Indonesia, education mismatch increases but is driven by the increase of undereducation. In contrast, the trend of overeducation decreased between 2000 and 2014 periods. The analysis is based on RM (Mode), by calculating the mode education level for a range of occupations with an individual defined as being overeducated if their qualifications are more than one standard deviation above their occupation's mode education level, and being undereducated if their qualifications are less than the standard deviation. In addition to the estimation techniques, the present study also develops a hybrid of occupation classifications to provide a good balance between a strong sample size and reducing the level of heterogeneity in roles within the occupational grouping and to provide a more accurate measurement of overeducated, undereducated and matched categories.

The result shows that the proportion of undereducated workers was 13.6 per cent in 2000, which then increased to 22.8 per cent in 2014. The overeducation proportions in 2000 and 2014 were around 28.5 per cent and 23.4 per cent, respectively. That conclusion is drawn based on both mode and mean methods. Thus, there is an increase of undereducation and a decrease of overeducation. Compared to OECD countries, undereducation and overeducation incidences in Indonesia are relatively higher. The average undereducation and overeducation incidences in OECD were 19 per cent and 15 per cent in 2015, respectively (World Bank, 2018).

One possible explanation for this trend is as explained by technological change theory (Oliveira *et al.*, 2000); rapid technological change may require school-provided skills/education higher than those possessed by the currently employed workers. Even in the presence of positive adjustment costs, better-educated workers cannot be made instantaneously. The employers and the employees could be locked into a situation of disequilibrium (at least in the short run) and hence pockets of undereducation would arise. Prospera and AlphaBeta Advisors (2019) find that rapid technological growth is occurring in Indonesia and transforming the nature of works in the country. Machines have eliminated around 5.5 hours of repetitive and menial tasks from the average Indonesian working week since 2000 and workers who were able to adapt to a higher value, more complex tasks saw their incomes grow almost twice as quickly over the period as those whose work did not automate. However, Indonesian firms have been slower to automate than their global peers; almost three out of ten are now investing in technology like car-assembly robots, agricultural tractors, and advanced point-of-sale systems. Another aspect that may attribute to the conclusion is related to the estimation techniques: the weakness of the RM is that it only considers the supply side, so the changes in the mode only reflect the changes in average workers' education. The measurement cannot represent the changes because the jobs demanded higher education qualifications or the change in education required by the firms.

In terms of the determinants of education mismatch, the present study develops a model of mismatch determinant using Multinomial Logit Model (MNL), since there is no standard model of mismatch determinants. Thus, most variables used by previous empirical studies (both internationally and specifically on Indonesia) are sorted, selected and adjusted based on the Indonesia context. The result shows that undereducation in Indonesia is determined by gender; there are some advantages for female workers compared to males, considering safety and respectability for example. The private sector experiences a higher relative probability of undereducation incident. This confirms the finding that the increase of undereducation incident is partly driven by the private sector. Moreover, undereducation is more likely to occur in non-agriculture sectors. Those working in medium and large firms are less likely to be undereducated, possibly because several large companies in the private sector have also developed their own training centres.

For overeducation, there is a negative relationship between tenure and the probability of being overeducated. The coefficient of the private variable is negative and significant, possibly because the Indonesian bureaucracy has a very rigid organisational structure in which the number of positions at each level in the hierarchy is fixed mechanically by a formula. Moreover, the probability of being overeducated is lower in manufacturing industries relative to the agriculture industry. Thus, the mismatch is determined by workers' and job's characteristics as the Assignment Model asserted.

Chapter 4 also analyses sector and gender differences with similar method (Multinomial Logit Model/MNL). Different trends of mismatch are found in the public sector; there is an increase in overeducation and a decrease in undereducation between 2000 and 2014. One of the possible reasons is the regulations. For instance, according to Government Regulation 30/2015, entry ranks are mainly determined by education level, and increases in rank are largely driven by seniority, with the maximum rank depending on the entry-level of the civil servant. More specifically, an undergraduate entry-level is IIIA and a master's is IIIB. In terms of wages, IIIB with 0-year experience receives a slightly higher wage (around GBP 7 per month) compared to IIIA with 0-year experience. Thus, workers with master's degrees still have an incentive (higher wages) to apply for the same position compared to those with undergraduate degree. Furthermore, the public sector recently prefers to hire workers with high education levels; those with at least senior high school qualifications. For unskilled jobs such as cleaning, however, the public sector prefers outsourcing or using private firms' services rather than hiring directly, as explained in Chapter 3. Nevertheless, there is no substantial difference between trends and the determinants of mismatch between males and females.

Several sensitivity tests are also performed to test the model, such as MNL with mean, Multinomial Probit (MNP) Model and by adding casual workers into the sample. The present study finds that the variables that determine educational mismatch are slightly sensitive to different methods, sets of variables, sector/gender and periods. Yet, the result still implies that educational mismatch in Indonesia is determined by personal and household characteristics, work-related and firm size and area of residency variables which are all observed in this research.

The last chapter, Chapter 5 finds, based on the Pooled OLS model, the wage return to an additional year of required schooling is 10.6 per cent, which is consistent with Duncan and Hoffman's (1981) finding. The wage returns to one year of surplus schooling is 10 per cent and the wage return to one year of deficit schooling is -8.7 per cent. The analysis employs Overeducation-Required-Undereducation (ORU) model, developed by Duncan and Hoffman (1981), to analyse the relationship between wage and education mismatch.

The model in the present study is relatively more systematic and comprehensive than the ones used in previous empirical studies since the model accommodates many variables that are grouped by personal characteristics, work-related, firm size and region categories. Meanwhile, panel data are used to deal with unobserved heterogeneity.

The test of equality of two coefficients shows that there is no significant difference between the wage of years of required schooling and years of surplus schooling. The implication here is still similar; the return to overeducation is positive and significant and the return to undereducation is negative and significant, both of which affect the wage. This finding is slightly different from previous empirical studies, such as Korpi and Tahlin (2007) and Tsai (2010). This could be related to Chapter 4's finding that overeducation trend decreases, implying that Indonesia needs more workers with higher education and better skill to meet the labour demand (Allen, 2016); as a result, overeducated workers will have at least the same wage as matched workers.

In comparison, the wage return to additional years of required schooling in the EU between 2006 and 2009 was 14 per cent; the wage return to one year of surplus schooling was 4 per cent; and the wage return to one year of deficit schooling was -3 per cent (Iriondo and Perez-Amaral, 2013). Meanwhile, the wage returns to additional years of required schooling in China between 1989 and 2009 was 23 per cent; the wage return to one year of surplus schooling was 3 per cent; and the wage return to one year of deficit schooling was -4 per cent. In other words, the wage return to an additional year of required schooling in Indonesia is relatively lower than in other countries. Also, the wage return to one year of surplus schooling is significantly higher and the wage return to one year of deficit schooling is significantly lower.

The panel data model is used to deal with unobserved heterogeneity. There are two methods of panel data: FE and RE. The crucial distinction between FE and RE is whether the unobserved individual effect embodies elements that are correlated with the regressors in the model, not whether these effects are stochastic or not. The main assumptions of RE model are random and uncorrelated with the independent variable included in the model. The advantage of FE model is that it can be used to analyse the impact of variables that vary over time, so the estimated coefficients of FE models cannot be biased because of the omitted time-invariant characteristics, such as culture, religion, gender, race, *etc*.

Turning to the results, RE result shows a conclusion similar to the pooled OLS model. Yet, applying FE substantially changes the coefficients value, pattern and significance level. The result is slightly different from previous empirical studies (Korpi and Tahlin, 2007; Tsai, 2010; Dockery and Miller, 2012) which find similar patterns between both models, although with lower coefficient values when using FE. The marked difference between the results of RE and FE may indicate that the ORU model in this study is relatively sensitive to the panel data methods used. Thus, a cautious interpretation is required. Another possible explanation is collinearity with the fixed effects, since the model accommodates several personal characteristic variables and year dummies which may be similar to individual level's fixed effects and the time's fixed effects. This is also in contrast to previous empirical studies which use fewer control variables. Alternatively, it could be argued that unobserved heterogeneity may affect the wage, though it is hard to prove.

From those three chapters' findings, there are some possible explanations for the inconsistent evidence with the previous empirical studies in the US and EU countries (Freeman, 1987; and McGuinness *et al.*, 2017). In the US, expansion in education is followed by the rise of overeducation incidences, since the supply of university graduates increases more rapidly than its demand (Freeman, 1976). The present study finds that Indonesia, as a developing country, has a different effect; the return to education tends to decline for most education levels between 2000 and 2014, and followed by the increase of undereducation. Some possible reasons are the different characteristics between the supply of and demand for labour. On the supply of labour, the proportion of labour force with higher education qualifications increases. However, some empirical studies find that the increase in education qualification is not followed by skill upgrading as low-quality education may occur (World Bank, 2018). This low quality is proven by the country's PISA score; Indonesia performs below the 25th percentile of OECD countries. Moreover, despite the education expansion, the proportion of population which have junior high school or lower qualifications is still relatively high (around 60 per cent of the total population) and the non-schooling population persists, accounting for around 4.5 per cent of the national population (aged 15 years or older) in 2016 (World Development Indicators, 2016). On the other hand, the demand for labour requires workers with high education qualifications as well as highly skilled workers; this is consistent with the economic transformation towards the service sector and the rapid technological change in the industries.

6.2 Implications

Several key findings from the present study are the returns to schooling in Indonesia tend to increase as the level of education increases; for trend overtime, the return to education declines while the education sector expands. By gender, the return to education is relatively higher for females than for males, and there is a combination effects of gender – years of schooling and the wages. By sector, the return to education in public sector is relatively higher than in private sector. The wage determination in public and private sector is different. Wages in private sector in Indonesia are largely determined by the market. On the other hand, the determination of public service's wage is more complicated, involving seniority, position, rank, and political approach. In addition, endogeneity problems do not occur, but there is a selection bias. This conclusion subject to an instruments' limitation problem in the present study.

In terms of trend in education mismatch and its determinants, the present study finds education mismatch increases between 2000 and 2014, as a consequence of the rise in under-education. The trend is similar when the sample is distinguished by gender. By sector, the private sector has a similar pattern with the main model. Meanwhile, the public sector has a contrary result; there is an increase in overeducation and a decrease in undereducation. In terms of determinants, the variables are slightly sensitive to the different methods used, the set of variables, the sector/gender and the periods. The results show that education mismatch in Indonesia is determined by personal and household characteristics, work related and firm size as well as area of residency variables which are all observed in this research and which is in line with the Assignment Models (mismatch is determined by workers' and job's characteristics).

In relations to the wages, both overeducation and undereducation incidences affect the labour wages in Indonesia. This finding is consistent with Duncan and Hoffman's (1981) finding, However, the test of equality of two coefficients shows that the difference between the coefficients of years of required schooling and surplus schooling is insignificant. Also, the penalty wages received by workers with years of deficit schooling is significantly different from the wage return to additional years of required schooling.

Thus, investing in higher education and improving the quality of the educational institutions are necessary for the government. There is an argument that a university qualification does not necessarily guarantee that individuals will fit the needs of the industries or workplaces, strengthening vocational education and skill training that can help prepare the youth for a smooth transition to employment; and promoting the integration of internship programs and other practical learning experience in the curriculums. Moreover, equal access to senior high schools and universities is necessary in terms of physical infrastructure (higher education building construction) as well as funding mechanisms (such as scholarship and student loan).

Those key findings contribute to the existing literature and future policies. For the literature, education expansion may affect the decline in the return to education which is consistent with an increase in the supply of educated labour (Walker and Zhu, 2008), but it is not always followed by an increase in overeducation incidences, as previous empirical literature commonly found (Freeman, 1987; and McGuinness *et al.,* 2017), or education expansion is followed by the decrease of both undereducation and overeducation (Yano, 2012). Furthermore, Indonesia has a different pattern, education expansion affects the decline in the return to education and is followed by the increase in undereducation incidences; possibly because labour demand lags behind labour supply.

Increasing education attainment, particularly senior high school and university levels, is required to meet the growing demand for workers with high skills/education in the labour force through 12-year compulsory programme, free education (like Germany or the UK), higher education scholarship, especially for students from low-income families. As Allen (2016) emphasises, Indonesia needs more educated and skilled workers to meet the labour demand. Yet, another possible consequence is more declining in the return to education. A similar situation happened in the US in 1970s (Freeman, 1977). Further demand-side adjustments could enhance the profitability of the investment in education and so increases the returns to their former levels. In the short run, the policymakers need to boost the labour demand through improving the investment climate, easing the starting of a business, and providing incentives for labour-intensive firms, particularly the low educated (as well as unskilled and low skilled) labours. It is also necessary to ensure that the speed of labour supply adjustment is as fast as the speed of labour demand and technology adjustment. When both the labour supply and demand are adjusted; reaching a better match between the labour supply and demand as well as a new equilibrium with higher wages. Another important implication is that information asymmetries should be minimised to enable a smoother matching process between job seekers and employers.

Another key finding of Chapter 3 is that the return to education in the public sector in Indonesia is much higher than in the private sector. The public sector continues to attract job seekers due to the appeals in the sector. This could create a crowding-out in the economy through many channels, such as the labour market where higher wages, more job security, or a higher probability of finding a public-sector job can make an individual more likely to seek or wait for a public-sector employment rather than searching for or accepting a job in the private sector. Also, individuals will only seek qualifications appropriate for entering the public sector through the education market rather than seeking the skills needed for productive employment in the private sector (Behar and Mok, 2013).

Education mismatch in Indonesia is determined by personal and household characteristics, work related and firm size as well as area of residency variables which are all observed in this research and which is in line with the Assignment Models (mismatch is determined by workers' and job's characteristics). The main role of the Ministry of Manpower is to minimise informational asymmetry. The Ministry of Manpower actually has a programme to facilitate matching of job seekers and job vacancies. The employment service gathers information on both vacancies and job seekers. It directs job seekers to employment which best suits their skills, ability and competencies. Any person has equal rights and opportunity to have access to labour market information, choose a job and earn a decent income to sustain decent living. But the government role is still not optimal since there is regional autonomy; thus, the job vacancy information and job seeker registration are conducted across the provinces and district level. Coordination, harmonisation and synchronisation rules and regulations are required. And the government must provide information on job opportunities through offline and online platforms.

Arguably, overeducation and undereducation may be related to job polarisation[72] in various ways. Job polarisation is closely linked to the shift from manufacturing to service sector, as a consequence of economic structural change; this also can occur due to rapid

[72] Job polarisation means that the share of high- and low-skill jobs grows at the expense of medium-skill jobs, and such trends have been linked to the decline in the demand for routine or codifiable tasks, including both manual and cognitive tasks.

technological change ILO (2017b). This leads to a change in the occupational structure and in skill/education-demand across jobs. For instance, if the growth of high-skill (education) occupations outpaces the supply of workers at this level of skills (education), undereducation can be expected to rise. On the other hand, overeducation may rise if high-skilled (educated) workers cannot find appropriate jobs and increasingly compete for jobs usually taken by those with a lower level of education (Sparreboom and Tarvid, 2016). ILO (2017) finds that the growth of the service industry in Indonesia was 7 per cent and the manufacturing industry was 4.4 per cent in 2016. Given that the largest industries are mining, construction and manufacturing, those industries require highly educated as well as skilled workers. Thus, the growth of high-skill (education) occupations may outpace the supply of workers at this level of skills or education in Indonesia; while the growth of low skill/education occupation may lag behind its supply, or abundant labour with lower education qualification (60 per cent of total labour) cannot be accommodated by the demand.

Furthermore, McGuinness *et al.* (2017) identify two types of undereducation. Firstly, undereducation and skill obsolescence; these are more likely to affect older workers and have similar drivers, such as technological change. And secondly, undereducation and skill gaps; these represent two approaches to describing the problem of deficit human capital among the workforce. Indonesia could have both types, undereducation associated with skill obstacle and skill gaps. Policies related to providing incentives to invest in training and to increase the number of training centres are required to be developed. Non formal education such as training and lifelong learning programme should be optimised and adjusted with labour market development, in particular for older generations, that experienced higher level of undereducation.

The declining return to education and increasing undereducation may highlight some inefficiencies in a country's educational system and labour market. Undereducation could affect weaker productivity growth and a slower structural transition to higher value-added activities with undereducated workforce (ILO, 2015). Communication between education providers and businesses is therefore essential to discover what is required by the businesses and how education providers can equip students with the necessary skills and knowledge. This can achieve a better alignment between educational/training supply and the labour demand. There should also be a strengthening of vocational education and skill training which can help to prepare the youth for a smooth transition to employment; and promoting of the integration between internship programs and other practical learning experiences in the education curriculum.

In terms of education mismatch and wage, the results indicate that workers with surplus year of education will have at least the same return with workers with required years of education, *ceteris paribus*. In the literature, overeducated workers will receive higher wages if they get a matching job. This may imply that one of the policies that can minimise education mismatch is unemployment insurance. Despite the existence of employment social security and health insurance programs, Indonesia still does not have an insurance scheme protecting the unemployed against poverty and assisting them before reemployment. World Economic Forum (2015) shows that there is evidence for positive effects of unemployment insurance on job quality; unemployment benefits can help workers to avoid large wage drops and even help them to join better firms. Meanwhile, Bosch (2016) finds that unemployment insurance can increase the formalisation of jobs; in turn, formal jobs may become more valued by workers, and it enables a more efficient job search. Thus, developing employment insurance scheme is urgently needed recently. In addition, the establishment of employment service centres also important, the main objectives are to provide training, vacancy information, and counselling for the unemployed as well as designing a system to monitor job search and participation in training activities.

6.3 Research Limitations and Extensions

Inevitably, there are limitations in this thesis. However, these limitations can be explored for better estimations in further research and used as a basis for the right policies.

Firstly, the interaction variables (between year dummies/gender and years of schooling) and wages significant affect the wages. Thus, it is worth to explore any interaction variables in the future research since the presence of interactions can have important implications for the interpretation of both Mincer wage equation and ORU model. If two variables of interest interact, the relationship between each of the interacting variables and a third "dependent variable" depends on the value of the other interacting variable.

Related to the issue of controlling for potential ability bias, the IV model could deal with this issue, if the instruments are appropriate. But IV is subject to an often-overlooked limitation as well. Many researchers intend to estimate the average treatment effect (ATE) for the entire population interest, but IV estimation only covers the local average treatment effect (LATE) or the ATE for the subpopulation that is influenced by the IV. When treatment effect are heterogenous across units, the LATE and the ATE may take on different values, this potentially causing complication in the interpretation, as well as makes comparison difficult, as explained in Chapter 3. Studies of the impact of over/under education on wages typically do not address this issue, because unobserved heterogeneity may influence education mismatch and its returns. The present study also finds that the result's interpretation is complicated since it is inconsistent with the idea of policy instruments and with the aim of the policies to increase the school participation rate, as explained in Section 3.4.7.

The present study does not consider school quality in the Mincer wage equation. Empirically, education differs between countries as it differs within countries, in terms of quality; for instance, individuals with the UK's university degree may receive higher wages than individuals with Indonesia's university degree, because education in the UK has better quality than education in Indonesia. As Card and Krueger (1992) find, men in the US in 1980 who were educated in states with higher-quality schools have a higher return to additional years of schooling. Thus, it is extremely necessary to measure the return to education by considering school quality, teacher's competence and other important factors to better understand how investment in education should be made. Purnastusti *et al*. (2015) also further that an analysis which considers education quality can give a clear understanding of factors affecting the return to education and can serve as an effective tool in the hands of organisations and institutions dealing with transition from school to work.

Furthermore, according to International Labour Organisation (2014), 53.6 per cent of the Indonesian labour force worked in informal employment. Rothenberg *et al.* (2015) emphasise that most of Indonesia's informal firms are very small, micro firms that pay low wages and are relatively unproductive. Thus, the return to education between the formal and informal employments will be different, as well as for waged and casual workers. However, the present study does not consider casual workers in the analysis since the wage data are only available for recent years. Further research on the return to education in casual workers or a comparison between the waged workers and casual workers would enrich the labour market analysis in Indonesia.

Related to job polarisation argument, the literature requires further analysis on the demand of labour based on occupations and skill levels. A more detailed analysis can map which different types of job (graduate/non-graduate jobs) and/or skilled/unskilled labours may have a substantial contribution to education mismatch. Meanwhile, the demand for labour can be analysed further using firm-level data or meta-analysis.

The present study also does not consider skill or at least subject study requirements since the data are not available in IFLS. Also, most companies will post job vacancies based on education and skill (or at least subject study), for instance: an automotive company requires workers with undergraduate degree in engineering for production staffs. Moreover, Badillo-Amador and Vila (2013) find that education and skill mismatches are two different phenomena of the labour market, although simultaneous. Specifically, for skill measurements, unfortunately the data are not available in IFLS and Indonesia Labour Force Survey (SAKERNAS). In the absence of skill and competence measurements, the results from this analysis should not be treated as conclusive, but rather as indicative that skill gap may affect the labour market productivity. There are some alternative measurements that can accommodate skill information, such as: (1) normative or professional job analysis (JA) - the method compares job titles with actual skill/education attainments based on information provided by professional job analysts; (2) subjective measures, including asking respondents directly or indirectly information on minimum job or skill requirements and their acquired education; or (3) mixed/alternative methods (EMX) or a combination between objective and subjective measures. Those measures can be elaborated in future research. Moreover, it is possible to combine the IFLS data with other data sources such as Village Potential Survey and National Socio-Economic Household Survey; that may provide data which is unavailable in the IFLS. There are some studies combines those data such as Wicaksono *et al,* (2018), the study examines the impact of parents' education and attending vocational high school to the probability of children attending tertiary education. Wicaksono *et al,* combined the IFLS and Village Potential Survey. Village Potential Survey provides the data about the education facility in location where children live, the data is in village level; while, the IFLS provides the data of the individual characteristics and household characteristics. One of advantages using PODES is more detail variable, such as: PODES provides data on school by institutions (public and private school), while IFLS only provide the number of schools in the villages.

The analysis can also be extended by taking into account skill mismatch and job satisfaction, as some studies (*e.g.* Badillo-Amador and Vila, 2013) argue that the consequences of education and skill mismatch is not only on wage, but also on job satisfaction. Badillo-Amador and Vila find that skill mismatches are perceived by workers as a much more relevant problem than education mismatches. The wage and job satisfaction consequences of skill mismatches are strongly negative; to the contrary, education mismatches show much weaker effects. Yet, Allen *et al.* (2001) argue that education mismatch and skill mismatch have their own advantages and disadvantages. Education mismatch affects wages strongly. Contrary to the assumption of assignment theory, this effect is not explained by skill mismatch. Conversely, skill mismatch is a much better predictor of job satisfaction and on-the-job search than education mismatch. Thus, extending the analysis could shed light on the mechanism of labour market adjustment or the dynamic of mismatch, since skill mismatch is an important cause of job dissatisfaction, and this provide an incentive for workers to look for other works, presumably one which is better suited to their abilities.

And finally, the IFLS data is longitudinal data that allows more deep analysis, beyond allowing for the modelling of unobserved heterogeneity, by looking at labour market dynamics such as: the dynamic of wages, in the sense that the analysis considers the persistency of wages: how long the employers and the employees could be locked into a situation of disequilibrium and whether education mismatch is temporary or a persistent phenomenon; overeducation or undereducation results in subsequent labour market transitioning behaviour. Thus, a clear understanding on education mismatch and wage relationship can inform the right policies, for example: unemployment benefits, since Indonesia has not had any policies on unemployment insurance yet. This is essential as such insurance can increase the formalisation of jobs; in turn, formal jobs may become more valued by workers, and it enables a more efficient job search (Bosch, 2016).

List of References

Abbasa, Q. and Foreman-Peck, J. (2007). *The Mincer Human Capital Model in Pakistan: Implications for Education Policy*. Cardiff business school working paper series, 44.

Abdulhafedh, A. (2017). Incorporating the Multinomial Logistic Regression in Vehicle Crash Severity Modeling: A Detailed Overview. *Journal of Transportation Technologies*, 7, 279-303. https://doi.org/10.4236/jtts.2017.73019

ACAPS (2016). *Data Cleaning*. [online] Acaps.org. Available at: https://www.acaps.org/sites/acaps/files/resources/files/acaps_technical_brief_data_cleaning_april_2016_0.pdf [Accessed 16 Sep. 2019].

Acemoglu, D. and Autor, D. (2011) Skills, Tasks and Technologies: Implications for Employment and Earnings. *Handbook of labor economics*, 4, pp. 1043–1171.

Acemoglu, D. (2014). *The Basic Theory of Human Capital*. Available at: https://econ.lse.ac.uk/staff/spischke/ec533/Acemoglu%20Autor%20chapter%201.pdf

Agresti, A. and Min, Y. (2002). *On Sample Size Guidelines for Teaching Inference about the Binomial Parameter in Introductory Statistics*. Unpublished Master's Thesis, University of Florida

AIPEG (2017). *Women's Economic Participation in Indonesia*. Available at: https://www.monash.edu/business/cdes/research/publications/publications2/Womens-economic-participation-in-Indonesia-June-2017.pdf.

Akresh, R., Halim, D., and Kleemans, M. (2018). *Long-term and Intergenerational Effects of Education: Evidence from School Construction in Indonesia.* Available at: https://delia-furtado.uconn.edu/wp-content/uploads/sites/1822/2019/03/Kleemans_schools.pdf

Alisjahbana, A. S., and Manning, C. (2006). Labour Market Dimensions of Poverty in Indonesia. *Bulletin Indonesian Economic Studies*, 42(2), 235-261. doi: 10.1080/00074910600873674.

Alisjahbana, A. S., Pitriyan, P. and Purnagunawan, R. M. (2017). *The Relevance or Irrelevance of Educational Mismatch for Science and Engineering Graduates in Indonesia.* AASLE.

Allen, J. P., Badillo-Amador, L., and van der Velden, R. K. W. (2013). *Wage Effects of Job-Worker Mismatches: Heterogeneous Skills or Institutional Effects?* GSBE Research Memorandum (071).

Allen, E. (2016). *Analysis of Trends and Challenges in the Indonesian Labor Market*. ADB Papers on Indonesia. Available at: https://www.adb.org/sites/default/files/publication/182935/ino-paper-16-2016.pdf.

Allen, E. R. and Kyloh, R. (2016). *Labour Reforms in Indonesia: An Agenda for Greater Equity and Efficiency.* Available at: https://www.ictu.ie/download/pdf/labour_reforms_in_indonesia_an_agenda_for_greater_equity_and_efficiency_november2016.pdf.

Amiti, M. and Cameron, L. (2011). *Trade Liberalization and the Wage Skill Premium: Evidence from Indonesia.* Available at: https://www.newyorkfed.org/medialibrary/media/research/economists/amiti/amiti_cameron_v54.pdf.

Anell, K. and Hartmann, D. (2007). *Flexible Work Arrangements in Asia. Boston College Center for Work and Family*. Available at: https://www.bc.edu/content/dam/files/centers/cwf/research/publications3/researchreports/Flexible%20Work%20Arrangements%20in%20Asia.

Angle, J. and Wissman, D. A. (1981). Gender, College Major, and Earnings. *Sociology of Education*, (54), pp. 25–33.

Angrist, J. D. and Krueger, A. B. (1991). Does Compulsory Schooling Attendance Affect Schooling Decisions. *Quarterly Journal of Economics*, 106(4), pp. 970-1014.

Antara (2017). *Menaker: Angkatan Kerja Indonesia hadapi masalah "mismatch*". Available at: https://www.antaranews.com/berita/652224/menaker-angkatan-kerja-indonesia-hadapi-masalah-mismatch.

Aronow, P.M. and Carnegie A. (2013). Beyond LATE: Estimation of the Average Treatment Effect with an Instrumental Variable. *Political Analysis*. Available at: https://www.jstor.org/stable/24572676#?seq=1

ASEAN Secretariat (2018). *Minimum Wage Levels Across ASEAN.* Available at: https://www.aseanbriefing.com/news/2018/08/30/minimum-wage-levels-across-asean.html.

Aslam, M., 2009. Education gender gaps in Pakistan: is the labor market to blame? *Economic Development Cultural Change*, 57 (4), 747–784

Asplund, R. and Pereira, P. T. (1999) *Returns to Human Capital in Europe: A Literature Review*.

Asplund, R. (2001). *Education and Earning: Further Evidence from Europe.* The Research Institute of the Finnish Economy.

Aswicahyono, H., Brooks, D. H. and Manning, C. (2011). *Exports and Employment in Indonesia: The Decline in Labor-Intensive Manufacturing and the Rise of Services*. ADB Economics Working Paper Series, 279(279), pp. 1–35. doi: 10.4324/9780203384824.x

Aswicahyono, H. and Manning C. (2016). *Labour Productivity, Wage and Management Issues in Labour-Intensive Industries.* AIPEG Report

Badillo-Amador, L. and Vila, L. E. (2013). Education and Skill Mismatches: Wage and Job Satisfaction Consequences. *International Journal of Manpower*, 34(5), pp. 416–428. doi: 10.1108/IJM-05-2013-0116.

Bahns, K. (2005). *Rural-to-Urban Migration in Developing Countries: The Applicability of the Harris Todaro Model with a Special Focus on the Chinese Economy*. Available at: https://pdfs.semanticscholar.org/90d7/598447b48f2bad76d427c5ea4d14477f6c7f.pdf.

Baker, M. J. and Jacobsen, J. P. (2007). Marriage, Specialization, and the Gender Division of Labor. *Journal of Labor Economics,* 25(4), pp. 763–793.

Baltagi, B. H. and Liu, L. (2015). Random Effects, Fixed Effects and Hausman's Test for the Generalized Mixed Regressive Spatial Autoregressive Panel Data Model. *Econometric Reviews*, 35(4), pp. 638-658.

Barro, Robert J. (1997). *Determinants of Economic Growth: A Cross-Country Empirical Study*. Cambridge MA, MIT Press.

Battu, H., Belfield, C.R., and Sloane, P.J. (1999). Overeducation among Graduates: A Cohort View. *Education Economics*, 7 (1), 21-38.

Battu, H. and Sloane, P. J. (2002). *Over-Education and Ethnic Minorities in Britain*. Manchester School, 72(4), pp. 535–559.

Bauer, T. K. (2002). Educational Mismatch and Wages: A Panel Analysis. *Economics of Education Review*, 21(3), 221-229.

Bauhoff S. (2014). Self-Report Bias in Estimating Cross-Sectional and Treatment Effects. In: Michalos A.C. (eds) Encyclopedia of Quality of Life and Well-Being Research. Springer, Dordrecht

Beatty, A., Berkhout, E., Bima, L., Coen, T., Pradhan, M. and Suryadarma, D. (2018). *Indonesian Children: In School but Not Learning*. Rise Programme in Indonesia. Available at: http://rise.smeru.or.id/sites/default/files/event/Daniel%20Suryadarma%20-%20Indonesian%20Children_In%20School%20but%20Not%20Learning.pdf.

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship. Available at SSRN: https://ssrn.com/abstract=1496221.

Becker, Gary S. (1971). *The Economics of Discrimination.* University of Chicago Press.

Becker, G. (1975). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. National Bureau of Economic Research.

Becker, G. S. (1994). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. 3rd Edition. The University of Chicago Press. Available at: https://www.nber.org/books/beck94-1.

Behar, A. and Mok, J. (2013). *Does Public-Sector Employment Fully Crowd Out Private-Sector Employment?* IMF Working Papers, 13(146). doi: 10.5089/9781484329412.001.

Baker, M. J. and Jacobsen, J. P. (2007). Marriage, Specialization, and the Gender Division of Labor. *Journal of Labor Economics,* 25(4), pp. 763–793. doi: 10.1086/522907.

Belfield. (2015). Weathering the Great Recession With Human Capital? Evidence on Labor Market Returns to Education From Arkansas. A CAPSEE Working Paper. Available at: https://files.eric.ed.gov/fulltext/ED562519.pdf.

Bell, A. J. D., & Jones, K. (2015). Explaining Fixed Effects: Random Effects Modelling of Time-Series Cross-Sectional and Panel Data. *Political Science Research and Methods*, 3(1), 133-153. https://doi.org/10.1017/psrm.2014.7

Belzil, C. and Hansen, J. (2002). Unobserved Ability and the Return to Schooling. *Econometrica*, 70, 575-591.

Belzil, C. (2006). *Testing the Specification of the Mincer Wage Equation*. Working Paper 06-08. Available at: hal.archives-ouvertes.fr/docs/00/14/25/42/PDF/0608.pdf.

Belzil, C. (2007). *Testing the Specification of the Mincer Wage Equation*. IZA Discussion Paper Series, (2650).

Berlingieri, F., Zierahn, U. (2014). *Field of Study, Qualification Mismatch, and Wages: Does sorting matter?* ZEW-Centre for European Economic Research Discussion Paper Nr.14-076.

Bhatti, S. H. (2012). *Estimation of the Mincerian Wage Model Addressing Its Specification and Different Econometric Issues.* General Mathematics. Université de Bourgogne.

Blackburn, M. L. and Neutmark, D. (1993). Are OLS Estimates of the Return to Schooling Biased Downward? Another Look. *Review of Economics and Statistics*, 77(2): 217-230.

Blundell, R., Dearden, L., Goodman, A. and Reed, H. (2000). The Returns to Higher Education in Britain: Evidence from a British Cohort. *The Economic Journal*, 110, pp.82-99.

Blundell, R., Dearden, L. and Sianesi, B. (2001). *Estimating the Returns to Education: Models, Methods and Results.* Centre for the Economics of Education, LSE, London.

Boll, C., Leppin, J., Rossen, A. and Wolf, A. (2016). *Overeducation - New Evidence for 25 European Countries*. IAB-Discussion Paper. Available at: http://doku.iab.de/discussionpapers/2016/dp3516.pdf.

Bosch, M. (2016). *Does Unemployment Insurance Offer Incentives to Take Jobs in the Formal Sector?* IZA World of Labor. doi: 10.15185/izawol.300.

British Council (2017). *Education Intelligence*. Available at: https://ei.britishcouncil.org/news/education-matters-indonesia.

Büchel, F., Pollmann-Schult, M. (2001). *Overeducation and skill endowments: the role of school achievement and vocational training quality*. IZA Discussion paper series No. 337, 1-28.

Büchel, F. and M. van Ham (2002). *Overeducation, Regional Labour Markets and Spatial Flexibility*. IZA Discussion Paper No. 424.

Büchel, F. and Pollmann-Schult, M. (2004). Overeducation and Human Capital Endowments. *International Journal of Manpower*, 25, pp. 150–166.

Bunzel, H. (2008). *Mincer Model*. [online] Available at: http://www2.econ.iastate.edu/classes/econ671/bunzel/Fall%202008/Mincerbeam.pdf [Accessed 16 Sep. 2019].

Burdett, K. and Coles, M. (2010). Wage/Tenure Contracts with Heterogeneous Firms. *Journal of Economic Theory,* 145(4), pp. 1408-1435.

Butcher, K. and Case, A. (1994). The Effect of Sibling Sex Composition on Women's Education and Earnings. *The Quarterly Journal of Economics,* 109(3), pp. 531-563.

Byron, R. P. and Takahashi, H. (1989). An Analysis of the Effect of Schooling, Experience and Sex on Earnings in the Government and Private Sectors of Urban Java. *Bulletin of Indonesian Economic Studies*, 25(1), pp. 105–117. doi: 10.1080/00074918812331335529.x

Cabinet Secretary (2018). *APBN 2018: Ada Angggaran BOS, Tunjangan Profesi Guru, dan Tunjangan Khusus Guru di Daerah Khusus.* Available at: https://setkab.go.id/apbn-2018-ada-anggaran-bos-tunjangan-profesi-guru-dan-tunjangan-khusus-guru-di-daerah-khusus/

Cameron and Trivedi (2010). *Microeconomics Using Stata*. Revised Edition.

Cameron and Miller (2015). *A Practitioner's Guide to Cluster-Robust Inference*.

Capsada-Munsech, Q. (2017). Overeducation: Concept, Theories, and Empirical Evidence. *Sociology Compass*, pp. 1–17. doi: 10.1111/soc4.12518.

Card, D. (1993). *Using Geographic Variation in College Proximity to Estimate the Return to Schooling*. NBER Working Paper No. 4483. doi: 10.3386/w4483.

Card, D. (1999). *The Causal Effect of Education on Earnings*. in Card, D. and Ashenfelter, O. (eds), Handbook of Labor Economics, Volume 3.

Card, D. (2001). Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. *Econometrica*, 69:5, pp1127-1160.

Card, D. and Krueger, A. B. (1992). Does school quality matter? Returns to education and the characteristics of public schools in the United States. *Journal of Political Economy*, 100(1), pp. 1–40.

Card, D. and Lemieux, T. (2001). Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis. *Quarterly Journal of Economics*, 116, May 2001, 705-746.

Carneval, A. and Rose, S. (2011). *The Undereducated American*. Available at: http://education.agu.org/files/2012/04/undereducated_american.pdf&embedded=true.

Cedefop (2009). *Skill Mismatch Identifying Priorities for Future*. Cedefop working paper. Available at: https://www.cedefop.europa.eu/files/6103_en.pdf.

Chua, K. and Chun, N. (2016). *In Search of a Better Match: Qualification Mismatches in Developing Asia*. ADB Economics Working Paper Series, (476). doi: http://dx.doi.org/10.2139/ssrn.2811502.

Clark, B. (2014). *The Career Prospects of Overeducated Americans The Career Prospects of Overeducated Americans.* IZA Discussion Paper Series, (8313).

Cohn, E. and Khan, S. P. (1995). The wage effects of overschooling revisited. *Labour Economics*, 2(1), pp. 67-76.

Comola, M. and de Mello, L. (2009). *The Determinants of Employment and Earnings in Indonesia: A Multinomial Selection Approach.* OECD, ECO/WKP(2009)31.

Comola, M. and de Mello, L. (2010). *Educational Attainment and Selection into the Labour Market: The Determinants of Employment and Warnings in Indonesia.* Available at: http://econpapers.repec.org/RePEc:hal:psewpa:halshs-00564835.

Comola, M. and de Mello, L. (2013). Salaried Employment and Earnings in Indonesia: New Evidence on the Selection Bias. *Applied Economics*, 45(19), pp. 2808–2816. doi: 10.1080/00036846.2012.667551.

Cedefop (2009). *Skill Mismatch Identifying Priorities for Future.* Cedefop working paper. Available at: https://www.cedefop.europa.eu/files/6103_en.pdf.

Cain, G, (1976). The Challenge of Segmented Labour Market Theories to Orthodox Theories. *Journal of Economic Literature*, 14:4, 1215-1257.

Chevalier, A. (2000). *Graduate Over-Education in the UK*.

Chevalier, A. (2003). Measuring Over-education. *Economica*. doi: 10.1111/1468-0335.t01-1-00296.

Chevalier, A. (2007). *Over education and skills of UK Graduates*. Centre for the Economics of Education, p. 42. doi: 10.1111/j.1467-985X.2008.00578.x.

Chevalier, A., Harmon, C. and Walker, I. (2002). *Does Education Raise Productivity, Reflect It ?*

Chevalier, A., Harmon, C., Walker, I. and Zhu, Y. (2004). Does Education Raise Productivity, or Just Reflect It? *The Economic Journal*, 114, pp.499-517.

Chevalier, A. and Lindley, J. (2009). Over Education and Skills of UK Graduates. Journal of the Royal Statistical Society: Series A. *Statistics in Society*, 172(2). doi: 10.1111/j.1467-985X.2008.00578.x

Chua, K. and Chun, N. (2016). *In Search of a Better Match: Qualification Mismatches in Developing Asia.* ADB Economics Working Paper Series, (476). Available at: https://www.adb.org/sites/default/files/publication/180978/ewp-476.pdf.

Clark, B., Joubert, C. and Maurel, A. (2012). *Overeducation and Skill Mismatch: A Dynamic Analysis.* Available at: https://www.semanticscholar.org/paper/Overeducation-and-skill-mismatch-%3A-a-dynamic-Clark-Joubert/b60d5fe9e6395632d61514a711802e8976c286b5#paper-header

Clark, B. (2014). *The Career Prospects of Overeducated Americans The Career Prospects of Overeducated Americans.* IZA DIscussion Paper Series, (8313).

Clark, B., Joubert, C. and Maurel, A. (2019). *The career prospects of overeducated Americans*. Available at: https://voxeu.org/article/career-prospects-overeducated-americans.

Cuevas, S., Mina, C., Barcenas, M., Rosario, A. (2009). *Informal Employment in Indonesia*. ADB Economics Working Paper Series, No. 156, Asian Development Bank (ADB), Manila, http://hdl.handle.net/11540/1805

Dasgupta, S., Bhula-Or, R. and Fakthong, T. (2015). *Earnings differentials between formal and informal employment in Thailand*. ILO Asia-Pacific Working Paper Series.

De Oliveira, M. M., Santos, M. C. and Kiker, B. F. (2000). The Role of Human Capital and Technological Change in Overeducation. *Economics of Education Review*, 19(2), pp. 199–206. doi: 10.1016/S0272-7757(99)00020-5.

Dhanani, S. and Islam, I. (2004). *Indonesian Wage Structure and Trends, 1976-2000.* Available at: http://www.ilo.int/public/english/protection/ses/download/docs/indonesia_wage.pdf.

Dickerson, A.P., (2005). *A Study on Rates of Return to Investment in Level 3 and Higher Qualifications*. Warwick Institute for Employment Research: University of Warwick.

Diem, A. (2015). Overeducation Among Graduates from Universities of Applied Sciences: Determinants and Consequences. *Journal of Economic & Financial Studies*, 03(02), pp. 63–77.

Di Gropello, E. and Sakellariou, C. (2010). *Industry and Skill Wage Premiums in East Asia*. The World Bank. Policy Research Working Paper, WPS 5379.

Di Gropello, E. and Kruse, A. and Tandon, P. (2011). *Skills for the Labor Market in Indonesia: Trend in Demand, Gaps, and Supply.* World Bank Publications, The World Bank, number 2282. doi: 10.1596/978-0-8213-8614-9.

Doan, T., Tuyen, T. Q. and Quan, L. (2016). *Lost in Transition? Declining Returns to Education in Vietnam*.

Dockery, A. M. and Miller, P. W. (2012). *Over-education, Under-education and Credentialism in the Australian Labour Market.* National Centre for Vocational Education Research (NCVER), Australian Government. Available at: https://files.eric.ed.gov/fulltext/ED535568.pdf.

Dolton, P. and Silles, M. (2002). *Determinants of Graduate Over-Education.* University of Oxford, Department of Economics Discussion Paper Series, (127). Available at: https://www.economics.ox.ac.uk/materials/working_papers/paper127.pdf

Dolton, P. and Vignoles, A. (2000). The Incidence and Effects of Overeducation in the U.K. Graduate Labour Market. *Economics of Education Review*, 19(2), pp. 179–198. doi: 10.1016/S0272-7757(97)00036-8.

Dong, S. X. (2016). Consistency between Sakernas and the IFLS for Analyses of Indonesia's Labour Market: A Cross-Validation Exercise. Bulletin of Indonesian Economic Studies, 52(3), pp. 343–378. doi: 10.1080/00074918.2016.1228828.

Dougherty, C. (2005). Why Are the Returns to Schooling Higher for Women than for Men? *The Journal of Human Resources*, 40(4), pp. 969-988.

Dow, J. K. and Endersby, J. W. (2004). Multinomial Probit and Multinomial Logit: A Comparison of Choice Models for Voting Research. *Electoral Studies*, 23, pp. 107–122. doi: 10.1016/S0261-3794(03)00040-4.

Duflo, E. (2001). Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *The American Economic Review*, 91(4), pp. 795–813. Available at: http://www.uh.edu/~adkugler/Duflo.pdf.

Duflo, E. (2002). *The Medium Run Effects of Educational Expansion Evidence from a Large School Construction Program in Indonesia*. Available at:

http://books.google.com/books?id=Up4WnQEACAAJ&dq=The+Medium+Run+Effects+of+Educational+Expansion&hl=&cd=1&source=gbs_api%5Cnpapers3://publication/uuid/71A83143-3767-4A8A-88DA-00B4F2E57167.

Dumauli, M. T. (2015). Estimate of the Private Return on Education in Indonesia: Evidence from Sibling Data. International Journal of Educational Development. *Elsevier Ltd*, 42, pp. 14–24. doi: 10.1016/j.ijedudev.2015.02.012.

Duncan, G. and Hoffman, S. (1981). The Incidence and Wage Effects of Overeducation. *Economics of Education Review*, 1(1), pp. 75-86.

Ekberg, J. and Salabasis, M. (2001). *The Firm Size Effect: Fact or Artifact?* SSE/EFI Working Paper Series in Economics and Finance, 462. Available at: https://opus.zbw-kiel.de/dspace/handle/10419/56112%5Cnhttp://swopec.hhs.se/hastef/papers/hastef0462.pdf.

Employment Policy Forum (2018). *Skill Development*. Policy Brief. Available at: https://laborforum-id.org/?information=fkk-10-policy-brief-skills-development.

Ertefaie, A., Small, D. S., Flory, J. H., and Hennessy, S. (2017). *A Tutorial on the Use of Innstrumenntal Variables in Pharmacoepidemiology.* Available at: https://www.cceb.med.upenn.edu/sites/default/files/uploads/cter/A%20tutorial%20on%20the%20use%20of%20instrumental_0.pdf

European Commission, Directorate-General for Employment, Social Affairs and Inclusion (2012). *Employment and Social Developments in Europe 2012*. Brussels.

Fahmi, M., Maulana, A. and Yusuf, A. A. (2011). *Teacher Certification in Indonesia: A Confusion of Means and Ends.* Working Paper in Economics and Development Studies Department, Department of Economics Padjadjaran University.

Farooq, A. (2016). *The U-shape of Over-education? Human Capital Dynamics and Occupational Mobility over the Lifecycle*. Available at: http://conference.iza.org/conference_files/empohe2016/farooq_a9620.pdf

Felipe, J. and Hasan, R. (2006). *Labor Markets in Asia*. New York: Palgrave Macmillan.

Fields, G. (2005). A Guide to Multisector Labour Market Models. Social Protection Discussion Paper Series. Available at: http://siteresources.worldbank.org/SOCIALPROTECTION/Resources/SP-Discussion-papers/Labor-Market-DP/0505.pdf

Fields, G. (2011). Labour market analysis for developing countries. *Labour Economics*. vol 18, pp. s16-s22.

Filmer, D. and Lindauer, D. L. (2001). Does Indonesia have a “low pay” civil service? Bulletin of Indonesian Economic Studies, 37(2), pp. 189–205. doi: 10.1080/00074910152390883.

Firpo, S., Fortin, N. and Lemieux, T. (2005). *Decomposing Wage Distributions: Estimation and Inference*. Available at: https://www.semanticscholar.org/paper/Decomposing-Wage-Distributions-%3A-Estimation-and-Firpo-PUC-Rio/1592e43029ca1175d89e3c537a4f274dc1576220

Flisi, S., Goglio, V., Meroni, E. C., Rodrigues, M., Vera-Toscano, E (2014). *Occupational mismatch in Europe: Understanding Overeducation and Overskilling for Policy Making.* European Commission, Joint Research Centre, Institute for the Protection and Security of the Citizen. doi: 10.2788/61733.

Flisi, S., Goglio, V., Meroni, E. C., Rodrigues, M., Vera-Toscano, E. (2017). Measuring Occupational Mismatch: Overeducation and Overskill in Europe—Evidence from PIAAC. *Social Indicators Research*. Springer Netherlands, 131(3), pp. 1211–1249. doi: 10.1007/s11205-016-1292-7.

Frank, R. H. (1978). Why Women Earn Less: The Theory and Estimation of Differential Overeducation. *American Economic Review,* 68 (3), 360–73.

Freeman, R. B. (1976). *The Overeducated American*. New York: Academic Press

Frenette, M. (2004). The overqualified Canadian Graduate: The Role of the Academic Program in the Incidence, Persistence, and Economic Returns to Overqualification. *Economics of Education Review*, 23(1), 29–45.

Foxton, F. (2016). *Analysis of the UK Labour Market - Estimates of Skills Mismatch Using Measures of Over and Undereducation.*

Furstenberg F. and Thrall C. (1975). Counting the Jobless: The Impact of Job-Rationing on the Measurement of Unemployment. *The Annals of the American Academy of Political and Social Science*, Vol. 418, No. 1, pp. 45-59.

Gautier, Pieter A., Teulings, Coen N. and Van Vuuren, Aico (2010). On-the-Job Search, Mismatch and Efficiency. *Review of Economic Studies*, Vol. 77, pp. 245-272.

Glewwe, P. (1996). The Relevance of Standard Estimates of Rates of Return to Schooling for Education Policy: A Critical Assessment. *Journal of Development Economics*, 51(2), pp. 267–290. doi: 10.1016/S0304-3878(96)00415-4.

Gosling, A. and Zhu, Y. (2010). *Overeducation*. GES Summer School, Kent.

Gottschalk, P. and Hansen, M. (2003). Is the Proportion of College Workers in Non- College College Jobs Increasing? *Journal of Labor Economics*, Vol. 21, pp. 449-471.

Green, F. and Henseke, G. (2016). *The Changing Graduate Labour Market: Analysis Using a New Indicator of Graduate Jobs.* IZA Journal of Labor Policy, 5(14). doi: 10.1186/s40173-016-0070-0.

Green, F., McIntosh, S. and Vignoles, A. (1999). *Overeducation and Skills Clarifying the Concepts*. CEP Discussion Papers dp0435, Centre for Economic Performance, LSE.

Green, F. and McIntosh, S. (2007). Is there a genuine under-utilisation of skills among the over-qualified? *Applied Economics,* 39, pp. 427-439.

Gronau, R. (1973). Wage Comparisons—A Selectivity Bias. *Journal of Political Economy*, 82(6), pp. 1119-1143

Groot, W. (1996). The Incidence of, and Returns to Overeducation in the UK. *Applied Economics*, 28, 1345-1350.

Groot, W. and Brink, H. M. Van Den (2000). Overeducation in the Labor Market: A Meta-Analysis. *Economics of Education Review*, 19, pp. 149–158. doi: 10.1086/298261.

Groot, W., Maassen Van Den Brink, H. (2003). *The Dynamics of Skill Mismatches in the Dutch Labour Market, Overeducation in Europe.* Current Issues in Theory and Policy. Edward Elgar Publishing, pp. 49–63

Grossman, M. (2008). The Relationship between Health and Schooling. *Eastern Economic Journal*, 34(3), pp. 281-292.

Guner, N., Kulikova, Y. and Llull, J. (2018). Marriage and Health: Selection, Protection, and Assortative Mating. *European Economic Review*, 104, pp. 138-166.

Handle, M. J. (2017). *Education and Skills Mismatch in Developing Countries: Magnitudes, Explanations, and Impacts.* Available at: https://www.ilo.org/wcmsp5/groups/public/---ed_emp/---ifp_skills/documents/presentation/wcms_554331.pdf.

Halaby C. N. (1994). Overeducation and Skill Mismatch. *Sociology of Education*, 67(1), pp. 47-59.

Harris, J.R. and M.P. Todaro (1970), Migration, Unemployment and Development: A Two-Sector Analysis. *American Economic Review*, 60, 126-142.

Harkness, S. and Machin, S. (1999). *Graduate Earnings in Britain 1974-1995*. DfEE Research Report 95.

Harmon, C., Oosterbeek, H. and Walker, I. (2000). *The Returns to Education A Review of Evidence, Issues and Deficiencies in the Literature.* Centre for the Economics of Education. London School of Economics and Political Science.

Harmon, C., Oosterbeek, H. and Walker, I. (2003). The Returns to Education: Microeconomics. *Journal of Economic Surveys*, 17(2). doi: https://doi.org/10.1111/1467-6419.00191.

Hartog, J., Oosterbeek, H., (1988). Education, allocation and earnings in the Netherlands: overschooling? *Economics Education Review*, 7(2), pp. 185-194.

Hartog, J. (1997). *On Returns to Education: Wandering Along the Hills of ORU Land*. Keynote speech for the LVIIth Conference of the Applied Econometrics Association, Maastricht.

Hartog, J. (2000). Over-education and earnings: where are we, where should we go? *Economics of Education Review*, 19, pp. 131–147.

Heckman, J. (1976). The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement*. 5 (4). pp. 475–492.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1), pp. 153-161.

Heckman, J. J., Lochner, L. J. and Todd, P. E. (2006). *Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond*. Available at: http://jenni.uchicago.edu/papers/pku_2007/Heckman_Lochner_etal_2006_HEE_v1_ch7.pdf

Heckman, J. J., Stixrud, J., Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labour Economics*, 24(3).

Hendajany, N. (2016). The Effectiveness of Public vs Private Schools in Indonesia. *Journal of Indonesian Applied Economics*, 6(1), pp.66-89.

Herlijanto, J. (2016). *What Does Indonesia's Pribumi Elite Think of Ethnic Chinese Today?* ISEAS Yusof Ishak Institute, 32(32), pp. 1–9.

HSBC (2017). *The Value of Education Higher and Higher*. Available at: https://www.hsbc.com/-/files/hsbc/media/media-release/2017/170628-the-value-of-education-higher-and-higher-global-report.pdf [Accessed 16 Sep. 2019].

Hu, Y., Scott, J. (2014). Family and Gender Values inn China: Generational, Geographic, and Gender Differences. *Journnal of Family Issues,* I-27. Available at: https://www.researchgate.net/publication/266202493_Family_and_Gender_Values_in_China_Generational_Geographic_and_Gender_Differences [Accessed 7 Feb. 2020].

Humphreys, J. (2012). *An Alternative to the Mincer Model of Education the Basic Mincer Model*. Available at: https://www.murdoch.edu.au/School-of-Business-and-Governance/_document/Australian-Conference-of-Economists/An-alternative-to-the-Mincer-model-of-education.pdf

ILO (2000). *Resolution concerning statistics of employment in the informal sector*, adopted by the Fifteenth International Conference of Labour Statisticians (January 1993); in: Current International Recommendations on Labour Statistics, 2000 Edition; International Labour Office, Geneva, 2000.

ILO (2002). *Decent work and the informal economy.* Available at: https://www.ilo.org/public/english/standards/relm/ilc/ilc90/pdf/rep-vi.pdf

ILO (2014). *Indonesia: Labour and Social Trends Update*. Asia-Pacific Decent Work Decade, pp. 1–4.

ILO (2017a). *Laporan Ketenagakerjaan Indonesia 2017: Memanfaatkan Teknologi untuk Pertumbuhan dan Penciptaan Lapangan Kerja.*

ILO (2017b). *Indonesia Jobs Outlook 2017: Harnessing technology for growth and job creation*. Available at: https://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---ilo-jakarta/documents/publication/wcms_613628.pdf.

ILO (2018). *Indonesia Continues to Improve Skills Development through Industry - Vocational Education and Training Partnership*. Available at: https://www.ilo.org/jakarta/info/public/pr/WCMS_654542/lang--en/index.htm.

Institute of Personnel and Development (1997). *Overqualified and Underemployed?* IPD Survey Report.

Inter-American Development Bank (2003). *Good Jobs Wanted*. (Washington: InterAmerican Development Bank).

Iriondo, I. and Pérez-Amaral, T. (2013). *The Effect of Educational Mismatch on Wages Using European Panel Data.*

Iriandoust, S. (2013). *Fulfilling Indonesia's Urgent Need to Develop Skilled Human Resources*. Available at: https://www.thejakartapost.com/news/2013/12/09/fulfilling-indonesia-s-urgent-need-develop-skilled-human-resources.html.

Jaeger, D. A. and Page, M. E. (1996). Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education. *The Review of Economics and Statistics*, 78(4), pp. 733–740.

Jones, M. K. and Slone, P. J. (2010). Disability and Skill Mismatch. *Economic Record*, 86(1), pp. 101-114.

JICA (2011). *Country Gender Profile: Indonesia*.

Kampelmann, S. and Rycx, F. (2012). The Impact of Educational Mismatch on Firm Productivity: Evidence from Linked Panel Data. *Economics of Education Review*, 31(6), pp. 918-931.

Katz, L. F., and Murphy, K. M. (1992). Changes in Relative Wages, 1963-1987: Supply and Demand Factors. *Ouvv Journal*, 107(February) 35-78.

Kiker, B. F., Santost, M. C. and de Oliveira, M. (1997). Overeducation and Undereducation: Evidence for Portugal. *Economics of Education Review*, 16(2), pp. 111–125.

Kim, D. and Topel, R. (1995). *Labor Markets and Economic Growth: Lessons from Korea's Industrialization, 1970-1990, Differences and Changes in Wage Structures*. Available at: http://www.nber.org/chapters/c7859.pdf.

Klaveren, M. Van et al. (2010). *An Overview of Women's Work and Employment in Indonesia Decisions for Life.* MDG3 Project', (14).

Kjelland, J. (2008). Economic Returns to Higher Education: Signaling v. Human Capital Theory; An Analysis of Competing Theories. *The Park Place Economist*, 16(1). [online] Available at: https://pdfs.semanticscholar.org/3049/c5a6d7dee397ebeda0666ad234584f23f60f.pdf

Kompas (2019). *Jokowi: Infrastruktur Bisa Ditinggal, Kita Fokus ke SDM*. Available at: https://nasional.kompas.com/read/2019/02/12/19541131/jokowi-infrastruktur-bisa-ditinggal-kita-fokus-ke-sdm.

Kontan (2018). *Pengangguran Lebih Banyak di Perkotaan*. Available at: https://nasional.kontan.co.id/news/pengangguran-lebih-banyak-di-perkotaan.

Korea-World Bank Partnership Facilities. (2015). The Indonesia Report on Standards and Qualifications (2015). Available at: http://documents.worldbank.org/curated/en/546521508745623197/pdf/120584-WP-P150980-PUBLIC-Indonesia-NQF-summary.pdf

Korpi, T. and Tåhlin, M. (2007). *Educational Mismatch, Wages, and Wage Growth: Overeducation in Sweden, 1974-2000.* Swedish Institute for Social Research (SOFI), 10.

Kuznets, S. (1955). Economic Growth and Income Inequality, *American Economic Review* 45: 1-28.

LaForge, G. (2016). *Changing a Civil Service Culture: Reforming Indonesia's Ministry of Finance, 2006 – 2010*. Innovation for Successful Societies, Princeton University. Available at: https://successfulsocieties.princeton.edu/sites/successfulsocieties/files/GL_CC_Indonesia_MOF_FORMATTED_5May2016_ToU.pdf

Law, E. H. and Miura, U (2015). *Transforming Teaching and Learning in Asia and the Pacific Case Studies from Seven Countries.* The United Nations Educational, Scientific and Cultural Organization and UNESCO Bangkok Office.

Lee, H., Lee, J. W. and Song, E. (2016). Effects of Educational Mistmatch on Wages in the Korean Labor Market. *Asian Economic Journal*, 30(4), pp. 375–400. doi: 10.1111/asej.12105.

Lee Kuan Yew School of Public Policy and Microsoft (2016). Technical and Vocational Education and Training in Indonesia: Challenges and Opportunities for the Future. Available at: https://lkyspp.nus.edu.sg/docs/default-source/case-studies/lkysppms_case_study__technical_and_vocational_education_and_training_in_indonesia.pdf?sfvrsn=e5c5960b_2

Lemieux T. (2006). *The "Mincer Equation" Thirty Years After Schooling, Experience, and Earnings*. In: Grossbard S. (eds) Jacob Mincer A Pioneer of Modern Labor Economics. Springer, Boston, MA.

Lewis, W. A. (1954). *Economic Development with Unlimited Supplies of Labour*. Manchester School, 22: 139-191.

Lewis, B. D. and Nguyen, H. T. M. (2018). *Policy Failure and Educational Attainment in Indonesia*. ANU Working Papers in Trade and Development, 17. Available at: https://acde.crawford.anu.edu.au/sites/default/files/publication/crawford01_cap_anu_edu_au/2018-08/final_2018_-_17.pdf

Li, H. and Urmanbetova, A. (2007). *The Effect of Education and Wage Determination in China's Rural Industry*. in Private Enterprises and China's Economic Development, edited by Shuanglin Lin and Xiaodong Zhu, 235-254. London: Routledge.

Lindley, J. and McIntosh, S. (2009). *A Panel Data Analysis of the Incidence and Impact of Over-education*. Working Paper. Department of Economics, University of Sheffield ISSN 1749-8368

Lindley, J. and Machin, S. (2016). The Rising Postgraduate Wage Premium. *Economica*, 83(330), pp. 281–306. doi: 10.1111/ecca.12184.

Linsley, I. (2005). Causes of Overeducation in the Australian Labour Market. *Australian Journal of Labour Economics,* 8(2), pp. 121–143.

Livanos, I. and Núñez, I. (2012). The Effect of Higher Education on the Gender Wage Gap. *International Journal of Education Economics and Development,* 3(1). doi: 10.1353/jda.2015.0072.

Livingstone, D. (1998). *The Education-Jobs Gap*. Westview Press, Boulder, CO, p. 341.

Lund Research (2013). *Measures of Central Tendency*. Laerd Stats. Available at: https://statistics.laerd.com/statistical-guides/measures-central-tendency-mean-mode-median.php.

Machin, S. and McNally, S. (2007). *Tertiary Education Systems and Labour Markets*. OECD. Available at: http://www.oecd.org/education/innovation-education/38006954.pdf

Manning, C. (2006). *A Review of Employment Trends and Statistics in Indonesia with Special Reference to the National Labour Force Survey (SAKERNAS)*. Draft report for the World Bank, Washington, DC.

Marimon, R. and Zilibbotti, F. (1999). Unemployment vs. Mismatch of Talents: Reconsidering Unemployment Benefits. *The Economic Journal,* 109(455), pp.266-291.

Mason, G. (1996). *Graduate Utilisation in British Industry: The Initial Impact Of Mass Higher Education*. National Institute Economic Review, May, pp.93-103.

Mavromaras, K., McGuinness, S. and Fok, Y.K. (2009). Assessing the Incidence and Wage Effects of Overskilling in the Australian Labour Market. *Economic Record*, 85(268), 60-72.

Mavromaras, K., Mcguinness, S., Leary, N. O., Sloane, P. (2019). *Job Mismatches and Labour Market Outcomes: Panel Evidence on University Graduates.* Available from: https://www.researchgate.net/publication/48264900_Job_Mismatches_and_Labour_Market_Outcomes_Panel_Evidence_on_University_Graduates.

McGoldrick, K. and Robst, J. (1996). *Gender Differences in Overeducation: A Test of the Theory of Differential Overqualification*. American Economic Review AEA Conference proceedings.

McGuinness, S. (2006). Overeducation in the Labour Market. *Journal of Economic Surveys*, 20(3), pp. 387–419.

McGuinness, S. & Sloane, P.J. (2011). Labour market mismatch among UK graduates: an analysis using REFLEX data. *Economics of Education Review*, 30: 130-145. http://dx.doi.org/10.1016/j.econedurev.2010.07.006

McGuinness, S., Bergin, A. and Whelan, A. (2017). *Overeducation in Europe: Trends, Convergence and Drivers*. IZA Institute Labor Economics. Available at: http://ftp.iza.org/dp10678.pdf

McGuinness, S., Pouliakas, K. and Redmond, P. (2017). *How useful is the concept of Skills Mismatch?* Available at: http://www.ilo.org/wcmsp5/groups/public/---ed_emp/---ifp_skills/documents/publication/wcms_552798.pdf.

McKinsey Global Institute (2012). *The Archipelago Economy: Unleashing Indonesia's Potential*. Available at: https://www.mckinsey.com/~/media/McKinsey/Featured%20Insights/Asia%20Pacific/The%20archipelago%20economy/MGI_Unleashing_Indonesia_potential_Full_report.ashx.

McLeod, R. H. (2006). Private Sector Lessons for Public Sector Reform in Indonesia. *Agenda: A Journal of Policy Analysis and Reform*, 13(3), pp. 275–288.

McLeod, R. H. and MacIntyre, A. (2007). *Democracy and the Promise of Good Governance*. Singapore, Institute of Southeast Asian Studies.

Mehta, A. et al. (2011). Overeducation in Developing Economies: How Can We Test For It, and What Does It Mean? *Economics of Education Review*, 30(6), pp. 1334–1347. doi: 10.1016/j.econedurev.2011.06.004.

Mercklé, P. and Octobre, S. (2015). *Do Survey Respondents Lie? Inconsistent Responses and the Biographical Illusion in a Longitudinal Survey on Adolescents' Leisure Outings*. Available at: https://www.cairn-int.info/article-E_RFS_563_0561--do-survey-respondents-lie.htm.

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution*. Journal of Political Economy,* 66(4), pp. 281-302.

Mincer, J. (1974). *Schooling, Experience and Earnings*. NBER. Cambridge MA.

Mincer, J. (1997). Changes in Wage Inequality, 1970-1990. *Research in Labor Economics*, 16, 1-18.

Ministry of Education and Culture (2015). *Indonesia Educational Statistics in Brief 2014/2015*.

Minister of Manpower (2009). *Roadmap Towards a Child Labour Free Indonesia in 2022.* Available at: https://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---ilo-jakarta/documents/publication/wcms_377170.pdf

Mishra, V. and Smyth, R. (2013). Economic Returns to Schooling for China's Korean Minority. *Journal of Asian Economics*, 24, pp. 89–102. doi: 10.1016/j.asieco.2012.09.001.

Moeliodihardjo, B. Y. (2013). *Higher Education Sector in Indonesia*. British Council.

Mohn, J. R. (2010). *Indonesian Snapshots in 2004, 2006, and 2008*. Available at: https://books.google.co.id/books?id=NIh0CQAAQBAJ&printsec=frontcover#v=onepage&q&f=false.

Montenegro, C. E. and Patrinos, H. A. (2012). *Returns to Schooling around the World*. Background paper for the World Development Report 2013. Available at: http://siteresources.worldbank.org/EXTNWDR2013/Resources/8258024-1320950747192/8260293-1320956712276/8261091-1348683883703/WDR2013_bp_Returns_to_Schooling_around_the_World.pdf

Moore, D., and McCabe, G. (1998). *Introduction to the Practice of Statistics*. 3rd ed. W. H. Freeman.

Morano, C. P. (2014). *The Determinants of Overeducation: Evidence from the Italian Labour Market*. Investigaciones de Economía de la Educación volume 9, in: Adela García Aracil & Isabel Neira Gómez (ed.), Investigaciones de Economía de la Educación 9, edition 1, volume 9, chapter 35, pages 681-698 Asociación de Economía de la Educación.

Moulton, Brent R. (1986). Random Group Effects and The Precision of Regression Estimates. *Journal of Econometrics*, 32(3):385–397.

Muffels, R. and Fouarge, D. (2001). *Working Profiles and Employment Regimes in European Panel Perspective*. OSA Institute for Labour Studies at Tilburg University Working Paper, 12.

Mućk, J. (2018). *Econometrics of Panel Data*. Available at: http://web.sgh.waw.pl/~jmuck/EoPD/Meeting1_2018.pdf.

Mullis, I., Martin, M., Foy, P. and Hooper, M. (2015). *TIMSS 2015 International Results in Mathematics*. Available at: http://timss2015.org/wp-content/uploads/filebase/full%20pdfs/T15-International-Results-in-Mathematics-Grade-4.pdf.

Mundlak, Y. (1978). Pooling of Time-series and Cross-section Data. *Econometrica*, 46(1):69–85

Muysken, J. and Weel, B., T. (1999). *Overeducation, Job Competition and Unemployment*. Available at: http://www.merit.unu.edu/publications/rmpdf/1999/rm1999-032.pdf [Accessed 16 Sep. 2019].

Nagib, L. and Ngadi (2008). Challenges of Unemployment in Indonenisa: Trends, Issues and Policies. Available at: http://ejurnal.kependudukan.lipi.go.id/index.php/jki/article/view/167

Nazara, S. and Safuan, S. (2005). Identifikasi Fenomena Overeducation di Pasar Tenaga Kerja Indonesia*. Jurnal Ekonomi Pembangunan Indonesia,* 6(1). doi: https://doi.org/10.21002/jepi.v6i1.152

Newhouse, D. and Beegle, K. (2006). The Effect of School Type on Academic Achievement: Evidence from Indonesia. *The Journal of Human Resources*, 41(3), pp. 529–557.

Nielsen, C. P. (2007). *Immigrant Overeducation: Evidence from Denmark*. World Bank Policy Research Working Paper No. 4234.

Nieto, S. and Ramos, R. (2017). *Overeducation, Skills and Wage Penalty: Evidence for Spain Using PIAAC Data.* Social Indicators Research, 134(1), pp. 219–236. doi: 10.1007/s11205-016-1423-1.

Nizam, N. and Nurdin, M. (2014). *Governance reforms in higher education: A study of institutional autonomy in Indonesia.* Available at: https://repository.ugm.ac.id/32577/1/UNESCO_RIHED_Study_-_Indonesia_(rev2).pdf

Nordin, M., Persson, I., & Rooth, D. (2010). Education–occupation mismatch: Is there an income penalty? *Economics of Education Review*, 29(6), 1047– 1059.

Nuffic (2017). *Education System Indonesia Described and Compared with the Dutch System*. Available at: https://www.nuffic.nl/documents/363/education-system-indonesia.pdf.

OECD (2017). *Pension at a Glace 2017: Country Profiles – Indonesia.* Available at: https://www.oecd.org/els/public-pensions/PAG2017-country-profile-Indonesia.pdf.

Ofek, H. and Y. Merrill. (1997). Labor mobility and the formation of gender wage gaps in local markets. *Economic Inquiry*, 35 (1): 28-47.

O'Leary, N.C. and Sloane, P.J. (2005). *The Changing Wage Return to an Undergraduate Education*. IZA Working Paper 1549

Ordine, P. and Rose, G. (2009). Overeducation and Instructional Quality: A Theoretical Model and Some Facts. *Journal of Human Capital*, 3(1), pp. 73–105. doi: 10.1086/599836.

Ortiz, L. (2010). Not the Right Job, but a Secure One Over-education and Temporary Employment in France, Italy and Spain. *Work, Employment and Society*, 24(1), 47–64. doi:10.1177/0950017009353657.

Park, H. M. (2011). *Regression Models for Panel Data Using Stata*. Available at: https://www.academia.edu/28505133/Practical_Guides_To_Panel_Data_Modeling_A_Step_by_Step.

Page, M. E. (2010). *Signaling in the Labor Market*. Elsevier. doi: 10.1016/B978-0-08-044894-7.01214-8.

Patrinos, H. (2016). *Estimating the Return to Schooling Using the Mincer Equation*. IZA World of Labor. doi: 10.15185/izawol.278.

Pecoraro, M. (2014). Is There Still a Wage Penalty for Being Overeducated but Well-Matched in Skills? A Panel Data Analysis of a Swiss Graduate. *Labour*, 28(3). Available at: https://doi.org/10.1111/labr.12031.

Pereira, P. T. and Martins, P. S. (2001). *Returns to Education and Wage Equations*. Institute for the Study of Labor (IZA), (298), pp. 1–37. doi: 10.1080/0003684042000217571.

Pissarides, C. A. (2000). *Equilibrium Unemployment Theory*. The MIT press.

Pietro, G. D. and Urwin, P. (2006). Education and Skills Mismatch in the Italian Graduate Labour Market. *Applied Economics*, 38(1), pp. 79–93. doi: 10.1080/00036840500215303.

Poteliene, S. and Tamasauskiene, Z. (2013). Human Capital Investment: Measuring Returns to Education. *Social Research*, 4(33), pp. 56–65.

Prospera and AlphaBeta Advisors (2019). *Capturing Indonesia's Automation Potential*.

Psacharopoulos, G. and Layard, R. (1979). Human Capital and Earnings: British Evidence and a Critique. *The Review of Economic Studies,* 46(3), p. 485. doi: 10.2307/2297015.

Psacharopoulos, G. (1994). Returns to Investment in Education: A Global Update. *World Development*, 22(9), pp. 1325–1343.

Psacharopoulos, G. and Patrinos, H. A. (2002). *Returns to Investment in Education: A Further Update*. World Bank Policy Research Working Papers 2881

Psacharopoulos, G. and Patrinos, H. A. (2004). Return to Investment in Education: A Further Update. *Education Economics*, 12(2), pp. 111–134. doi: 10.1080/0964529042000239140.

Purnastuti, L., Miller, P. W. and Salim, R. (2013). Declining Rates of Return to Education: Evidence for Indonesia. *Bulletin of Indonesian Economic Studies*, 49(2), pp. 213–236. doi: 10.1080/00074918.2013.809842.

Purnastuti, L., Salim, R. and Joarder, M. A. M. (2015). The Returns to Education in Indonesia: Post Reform Estimates. *The Journal of Developing Areas*, 49(3), pp. 183–204. doi: 10.1353/jda.2015.0174.

Quang, H.L., and Tran-Nam, B. (2019). Qualification mismatch in the labour market and the impact on earnings: evidence from Vietnam. *Journal Economics and Development*. Available at: https://www.emerald.com/insight/content/doi/10.1108/JED-09-2019-0032/full/html

Quintini G. (2011). *Right for the Job: Over-Qualified or Under-Skilled?* OECD, Directorate for Employment, Labour and Social Affairs, Working Document. SEM (2011), 5.

Rani P G (2014). Disparities in Earnings and Education in India. *Cogent Economics and Finance*, 2: 941510. Retrieved from http://dx.doi.org/10.1080/23322039.2014.941510.

Rao, S. (2015). *Is the Private Sector More Efficient? A Cautionary Tale?* UNDP Global Centre for Public Service Excellence. Available at: https://www.undp.org/content/undp/en/home/librarypage/capacity-building/global-centre-for-public-service-excellence/efficiency.html

Regan, T. L. and Oaxaca, R. L. (2009). Work Experience as a Source of Specification Error in Earnings Models: Implications for Gender Wage Decompositions. *Journal of Population Economics*, 22(2), pp. 463-499.

Ren, W. and Miller, P. W. (2012). Gender Differentials in the Payoff to Schooling in Rural China. *The Journal of Development Studies,* 48(1), pp. 133-150.

Rigg, M., Elias, P., White, M. and Johnson, S. (1990). *An Overview of the Demand for Graduates*. HMSO, London.

Reis, M. (2015) *Overeducation, undereducation and labor earnings in Brazil : a panel data analysis*.

Rothenberg, A. D. et al. (2016). Rethinking Indonesia's Informal Sector. *World Development*, 80, pp. 96–113. doi: 10.1016/j.worlddev.2015.11.005.

Rodriguez, S. B. (2011). Are Educational Mismatches Responsible for the 'Inequality Increasing Effect' of Education? *Social Indicators Research*, 102(3), pp. 409-437. DOI: 10.1007/s11205-010-9675-7

Rumberger, Russell W. (1981). *Overeducation in the U.S. Labor Market*. Praeger Publishers: New York.

Rumberger, R. (1987). The Impact of Surplus Education on Productivity and Earnings. *Journal of Human Resources*, 22(1), pp. 24–50.

Ryznar, R. and Vidican, G. (2006). *Quantitative Reasoning and Statistical Methods for Planning*. Available at: http://web.mit.edu/11.220/www05/lab3_06/regression1.htm.

Sakinah, N. (2017). Employees' Attitude toward the Implementation of Open Selection System for Senior Executive Service Positions. *International Journal of Administrative Science & Organization.* Vol 24, No. 1.

Saputra, M. (2018). *Who Wants to Be an Indonesian Civil Servant? Almost Everyone*. Available at: https://thediplomat.com/2018/10/who-wants-to-be-an-indonesian-civil-servant-almost-everyone/ [Accessed 16 Sep. 2019].

Sattinger, M. (1993). Assignment models of the distribution of earnings. *Journal of Economic Literature*, 31(2), pp. 831–880.

Sattinger, M. (2012). *Assignment Models and Quantitative Mismatches*. Available at: https://www.albany.edu/economics/research/seminar/files/Michael%20Sattinger.pdf

Schaner, S. and Das, S. (2016). *Female Labor Force Participation in Asia: Indonesia Country Study*. ADB Economics Working Paper Series, (474), pp. 1–50.

Schultz, T. W. (1961). Investment in Human Capital. *The American Economic Review*, 51(1), pp. 1-17.

Schurer, S. and Jongsay Y. (2012). *Personality, Well-being and the Marginal Utility of Income: What Can We Learn from Random Coefficient Models?* Health, Economics and Data Group. Working Paper, University of York. Available at http://www.york.ac.uk/res/herc/documents/wp/12_01.pdf.

Sharma, A. (2013). *Informality, Firm Size and Labour Market Segmentation in Indonesia*. Available at: https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=CSAE2014&paper_id=771

Sicherman, N. (1991). Overeducation in the Labor Market. *Journal of Labor Economics*, 9(2), pp. 101–122.

Sicherman, N., & Galor, O. (1990). A theory of career mobility. *Journal of Political Economy*, 98(1), pp. 169–192.

Sindo (2018). *3,6 Juta Pendaftar Berebut untuk 238.015 Kursi CPNS*. Available at: https://jatim.sindonews.com/read/2418/1/36-juta-pendaftar-berebut-untuk-238015-kursi-cpns-1539767421.

Silles, M. and Dolton, P. (2002). *The Determinants of Graduate Over-Education.* University of Oxford, Department of Economics Discussion Paper Series, (127).

Singhari, S. and Madheswaran, S. (2016). *The Changing Rates of Return to Education in India: Evidence from NSS Data.*

Sloane, P. J., Battu, H. and Seaman, P. T. (1999). Overeducation, Undereducation and the British Labour Market. *Applied Economics*, 31(11), pp. 1437–1453. doi: 10.1080/000368499323319.

SMERU (2002). *Minimum Wages: Social Policy versus Economic Policy*. SMERU Newsletter, 01. Available at: http://www.smeru.or.id/sites/default/files/publication/news01.pdf.

Sparreboom, T. and Tarvid, A. (2016). Imbalanced Job Polarization and Skills Mismatch in Europe. *Journal for Labour Market Research*, 49, pp. 15–42. doi: 10.1007/s12651-016-0196-y.

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, 87(3), pp. 355–374. doi: 10.2307/1882010.

Strauss, J., Witoelar, F., and Sikoki, B. (2016). The Fifth Wave of the Indonesia Family Life Survey: Overview and Field Report. Available at: https://www.rand.org/content/dam/rand/pubs/working_papers/WR1100/WR1143z1/RAND_WR1143z1.pdf

Steiner, V. and Wagner, K. W. (1996). *Has earnings inequality in Germany changed in the 1980's?*

Starkweather, J. and Moske, A. K. (2011). *Multinomial logistic regression*. Available at: https://it.unt.edu/sites/default/files/mlr_jds_aug2011.pdf.

Statistics Indonesia (2019). *Education Statistics*. Available at: https://www.bps.go.id/subject/28/pendidikan.html#subjekViewTab3.

Statistics Indonesia and ILO (2010). *Working Children in Indonesia 2009*. Available at: https://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---ilo-jakarta/documents/publication/wcms_123585.pdf.

Steiner, Viktor; Wagner, Kersten (1996). *Has earnings inequality in Germany changed in the 1980's?* ZEW Discussion Papers, No. 96-32, Zentrum für Europäische Wirtschaftsforschung (ZEW), Mannheim.

Suharti (2013). *Trends in Education Indonesia*, in Suryadarma, D. and Jones, G.W. (eds) Education in Indonesia. Singapore: ISEAS.

Suryahadi, A., Widyanti, W., Perwira, D., and Sumarto S. (2003). Minimum Wage Policy and Its Impact on Employment in the Urban Formal Sector. *Bulletin of Indonesian Economic Studies*, 39(1), pp. 29-50

Suryadarma, D., Suryahadi, A., Sumarto, S. and Rogers, F. (2004). *The Determinants of Student Performance in Indonesian Public Primary Schools: The Role of Teachers and Schools.* Available at: https://www.smeru.or.id/sites/default/files/publication/determinanteng.pdf [Accessed 16 Sep. 2019].

Suryadarma, D., Widyanti, W., Suryahadi, A., and Sumarto, S. (2006). *From Access to Income: Regional and Ethnic Inequality in Indonesia*. SMERU Working Paper. Available at: http://www.smeru.or.id/sites/default/files/publication/fromaccesstoincome_eng.pdf

Swaffield, J. (2000). *Gender, Motivation, Experience and Wages*. Available at: http://eprints.lse.ac.uk/20188/1/Gender%2C_Motivation%2C_Experience_and_Wages.pdf.

Taniguchi, K., & Tuwo, A. (2014). *New evidence on the gender wage gap in Indonesia*. Manila: Asian Development Bank.

Tarvid, A. (2015). The Role of Industry in the Prevalence of Overeducation in Europe. *Procedia Economics and Finance, Elsevier B.V.*, 30(15), pp. 876–884. doi: 10.1016/S2212-5671(15)01337-4.

Teixeira, P. N. (2014). *Gary Becker's Early Work on Human Capital – Collaborations and Distinctiveness*. IZA Journal of Labor Economics, 3(1), pp. 1–20. doi: 10.1186/s40172-014-0012-2.

The Boston Consulting Group (2013). *Growing Pains, Lasting Advantage: Tackling Indonesia's Talent Challenges*. Available at: http://image-src.bcg.com/Images/Growing_Pains_Lasting_Advantage_May_2013_tcm9-97662.pdf.

The Jakarta Post (2016). *Indonesia Seen to Have 1 Million Excess Civil Servants*. Available at: https://www.thejakartapost.com/news/2016/06/07/indonesia-seen-to-have-1-million-excess-civil-servants.html [Accessed 16 Sep. 2019].

The Jakarta Post (2018). *Commentary: Are we heading toward demographic bonus or disaster?* Available at: https://www.thejakartapost.com/academia/2018/03/05/commentary-are-we-heading-toward-demographic-bonus-or-disaster.html.

Thurow, L. C. (1975). *Generating inequality: Mechanisms of distribution in the US economy*. Basic books.

Tinbergen, J. (1951). Some remarks on the distribution of labour incomes. *International Economic Papers*, 1, pp. 195–207.

Tjahjono, T. (2017). *Why Civil Servant is Highly Attractive Job in Indonesia*. Available at: https://www.globalindonesianvoices.com/31125/why-civil-servant-is-highly-attractive-job-in-indonesia/.

Tjiptoherijanto, P. (2014). *Reform of the Indonesian Civil Service: Racing with Decentralization*.

Tjiptoherijanto, P. (2018). *Reform of the Indonesian Civil Service: Looking for Quality*. Economics World, Nov.-Dec. 2018, Vol. 6, No. 6, 433-443. Available at: https://www.davidpublisher.org/Public/uploads/Contribute/5c25935a8174a.pdf

TNP2K. (2015). *Supply of Non-Formal Training in Indonesia*. Available at: http://www.tnp2k.go.id/images/uploads/downloads/WP_23-1.pdf

Tobias, J., Wales, J., Syamsulhakim, E. and Suharti (2014). *Towards Better Education Quality: Indonesia's Promising Path*. Available at: https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9065.pdf.

Topel, R. (1991). Specific Capital, Mobility and Wages: Wages rise with Job Seniority", *Journal of Political Economy*, 99(1), pp. 145-176.

Triola, M. F. (2000). *Elementary Statistics*. 8th ed. Addison-Wesley.

Trostel, P., Walker, I. and Woolley, P. (2002). Estimates of the Economic Return to Schooling for 28 Countries. *Labour Economics*, 9, pp. 1–16.

Tsai, Y. (2010). Economics of Education Review Returns to Overeducation: A longitudinal analysis of the U.S. labor Market. *Economics of Education Review*. Elsevier Ltd, 29(4), pp. 606–617. doi: 10.1016/j.econedurev.2010.01.001.

Tsang M. C. and Levin H. M. (1985). The Economics of Overeducation. *Economics of Education Review*, 4(2), pp. 93-104.

Tsang, M. (1987). The Impact of Underutilisation of Education of Productivity: A Case Study of the U.S. Bell Companies. *Economics of Education Review,* 11, pp. 239–234

Tsang, M., Rumberger, R. and Levin, H. (1991). The Impact of Surplus Schooling on Worker Productivity. *Industrial Relations*, 30(2), pp. 209–228.

Tyler, J. H., Murnane, R. J. and Willett, J. B. (2000). Estimating the Labor Market Signaling Value of the GED. *The Quarterly Journal of Economics*, 115(2), pp. 431–468. doi: 10.1162/003355300554818.

UNDP (2014). *Motivation of Public Service Officials, Insights for Practitioners*. UNDP Global Centre for Public Service Excellence report, (3), pp. 2011–2014. Available at: http://www.undp.org/content/undp/en/home/librarypage/capacity-building/global-centre-for-public-service-excellence/motivation.html.

UNDP. (2014). *Motivation of Public Service Officials, Insights for Practitioners*. UNDP Global Centre for Public Service Excellence report, (3), pp. 2011–2014. Available at: http://www.undp.org/content/undp/en/home/librarypage/capacity-building/global-centre-for-public-service-excellence/motivation.html.

UNDP Global Centre for Public Service Excellence (2015). *Is the Private Sector More Efficient? A Cautionary Tale.* Available at: https://www.undp.org/content/dam/undp/library/capacity-development/English/Singapore%20Centre/GCPSE_Efficiency.pdf.

UNESCO (2010). *Measuring Educational Participation*, 6(4), pp. 1–20.

UNESCO (2019). *Gross Enrolment Ratio*. Available at: http://uis.unesco.org/en/glossary-term/gross-enrolment-ratio.

Verdugo, R. R. and Verdugo, N. T. (1989). The Impact of Surplus Schooling on Earnings: Some Additional Findings. *Journal of Human Resources*, 24(4), pp. 629-643.

Wachter, M. (1974). *Primary and Secondary Labour Markets: A Critique of the Dual Approach*. Brookings Papers on Economic Activity, 3: 637–680.

Wajdi, N., Mulder, C. H. and Adioetomo, S. M. (2017). Inter-regional Migration in Indonesia: A Micro Approach. *Journal of Population Research*, Springer Netherlands, 34(3), pp. 253–277. doi: 10.1007/s12546-017-9191-6.

Walker, I. and Zhu, Y. (2003). *Education, Earnings and Productivity*.

Walker, I. and Zhu, Y. (2005). *The college wage premium, overeducation and the expansion of higher education in the UK.* IZA DP No. 1627.

Walker, I. and Zhu, Y. (2008). The College Wage Premium and the Expansion of Higher Education in the UK. *Scandinavian Journal of Economics*, 110(4), pp. 695–709. doi: 10.1111/j.1467-9442.2008.00557.x

Wang, H. (1995). *Do Marital Status and Presence of Children Affect Women's Wage Rate?* Consumer Interest Annual, 41.

Warner, R.M. (2013). Applied Statistics from Bivariate Through Multivariate Techniques. Sage.

Weber, A. (2019). *Do Unemployment Benefits Help People Find Better Jobs?* World Economic Forum. Available at: https://www.weforum.org/agenda/2015/07/do-unemployment-benefits-help-people-find-better-jobs/.

Weiss, A. (1995). Human Capital vs. Signaling Explanations of Wages. *Journal of Economic Perspectives*, 9(4), pp. 133–154.

WEF (2015). *The Human Capital Report 2015*. World Economic Forum. doi: www3.weforum.org/docs/WEF_Human_Capital_Report_2015.pdf.

Wen, L. and Maani, S. A. (2017). *A Panel Study of Immigrants' Overeducation and Earnings in Australia*. IZA Institute of Labor Economics. Available at: http://ftp.iza.org/dp11216.pdf

WENR (2019). *Education in Indonesia*. WENR World Education News and Reviews. Available at: https://wenr.wes.org/2019/03/education-in-indonesia-2 [Accessed 19 Sep. 2019].

Wicaksono, W. Sparrow, R. A. and van Bergeijk, P. (2018). The Impact of Parents Education and Attending Vocational High School to College Entrance. *Jurnal Pendidikan Vokasi*. Vol. 8, No.1, pp. 12-23

Wiggins, S. and Keats, S. (2014*). Rural Wages in Asia*. Available at: http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9186.pdf.

Wihardja (2016). *Indonesia's Structural Transformation Offers Clues on Where to Find Good Jobs*. Available at: http://blogs.worldbank.org/eastasiapacific/indonesia-s-structural-transformation-offers-clues-where-find-good-jobs on 11 June 2019\

Wooldridge, J.M. (2009). Introductory Econometrics A Modern Approach. South-Western Cengage Learning, USA.

World Bank (2000). *Administrative and Civil Service Reform*. Available at: http://www1.worldbank.org/publicsector/civilservice/countries/indonesia/shapesize.htm.

World Bank (2010). *Indonesia Skills Report: Trends in Skills Demand, Gaps, and Supply in Indonesia*.

World Bank (2011). *World Development Report: Gender Equality and Development*. World Development Report 2012. doi: 10.1596/9780821388105_ch5.

World Bank (2013). *Spending more or spending better: Improving education financing in Indonesia.* Available at: http://documents.worldbank.org/curated/en/991011468254684449/pdf/730500ESW0Whit0ain0report0March2013.pdf

World Bank. (2014). *Indonesia: Avoiding the Trap*. Available at: https://www.worldbank.org/content/dam/Worldbank/document/EAP/Indonesia/Indonesia-development-policy-review-2014-english.pdf

World Bank. (2015). Indonesia - National Qualifications Framework summary (English). Washington, D.C. : World Bank Group. http://documents.worldbank.org/curated/en/546521508745623197/Indonesia-National-Qualifications-Framework-summary

World Bank (2016). *World Development Report 2016: Digital Dividends*. Available at: https://www.worldbank.org/en/publication/wdr2016.

World Bank (2018). *World Development Report 2018: Learning to Realize Education's Promise*. Washington.

World Bank. (2019). *World Development Indicators*. World Bank. Available at: http://datatopics.worldbank.org/world-development-indicators/

World Economic Forum (2015). *Do Unemployment Benefits Help People Find Better Jobs?* Available at: https://www.weforum.org/agenda/2015/07/do-unemployment-benefits-help-people-find-better-jobs/

World Economic Forum (2015). *The Human Capital Report 2015*. World Economic Forum. doi: www3.weforum.org/docs/WEF_Human_Capital_Report_2015.pdf.

Wye, C. and Ismail, R. (2018). The Effects of Ability on Returns to Over- and Under-Education: Evidence from Malaysia. *International Labour Review*. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/ilr.12094.

Yano, S. (2012). *Overeducated? The Impact of Higher Education Expansion in Post-Transition Mongolia*. Columbia University. Available at: http://search.proquest.com/docview/1014166542?accountid=41453.

Yin, L. (2016). *Overeducation in the Chinese Labour Ma*rket. PhD Thesis, The University of Sheffield.

Appendices

Appendix I: IFLS Survey

Figure I.1: IFLS1 Sampling
Source: IFLS1 User Guide.
Note: IFLS1 was conducted in grey areas.

Table I.1: 1993 Dynasty Re-Contact Rates

	IFLS1	IFLS2		IFLS3		IFLS4		IFLS5		All rounds	
	HH	HH	Re-contact rates (%)	HH	Re-contact rates (%)	HH	Re-contact rates (%)	HH	Re-contact rates (%)	HH	Re-contact rates (%)
Dynasty contacted	7,224	6,821	94.4	6,883	95.3	6,761	93.6	6,647	92.0	6,341	87.8
Dynasty interviewed		6,752	93.5	6,787	94	6,553	90.7	6,555	90.7	6,275	86.9
Dynasty died (cumulative)		69	1	97	1.3	211	2.9	317	4.4	-	-
Dynasty not contacted		403	5.6	341	4.7	463	6.4	577	8.0	-	-

Source: IFLS2, IFLS3, IFLS4, IFLS5
Number of dynasties contacted includes those whose members all died and households that recombined into other households since the last survey.

Source: IFLS5 User Guide.

Figure I.2: Households Re-Contact Rates IFLS1-IFLS5

	IFLS1	IFLS2 target households contacted	Re-contact rate (%)	IFLS3 target households	IFLS3 target households contacted	Re-contact rate (%)	IFLS4 target households	IFLS4 Target Contacted	Re-contact rate (%)	IFLS5 target households	IFLS5 Target Contacted	Re-contact rate (%)
IFLS1households	7,224	6,821	94.42	7,138	6,800	95.3	7,135	6,596	92.4	7,131	6,432	90.2
IFLS2 split-off households	-	877	-	865	819	94.7	876	769	87.8	703	650	92.5
IFLS2+ splitoff households	-	-	-	344	309	89.8	335	295	88.1	243	224	92.2
IFLS3 splitoff households	-	-	-		2,646		2,648	2,302	86.9	2,164	1,923	88.9
IFLS3 target households				8,347		95						
IFLS4 splitoff households								4,033		4033	3687	91.4
IFLS4 target households	-	-	-				10,994	9,962	90.6			
IFLS5 target households										14,274	12,916	90.5
IFLS5 splitoff households	-										4,015	
Total households	7,224	7,698			10,574			13,995			16,931	

Source: IFLS2, IFLS3, IFLS4, IFLS5
Number of households contacted includes those whose members all died and households that recombined into other households since the last survey.

Source: IFLS5 User Guide.

Table I.3: Households Sample and Completion Rate, IFLS1-IFLS5

		1990 Population			IFLS 1 HH Ivwd	IFLS2 Households: Interviewed, died, or joined other hh					IFLS3 Households: Interviewed, died, or joined other hh				
	Provinces [a]	N(000)[b]	%	IFLS EAs		% IFLS 1 HH	# IFLS1 HH	Any split-off HH	Total	Inter-viewed	% IFLS1 HH	# IFLS1 HH	Any split-off HH	Total	Inter-viewed
11	Aceh	3,476	1.9												
12	**North Sumatra**	10,391	5.7	26	563	89.5	504	44	548	545	90.7	507	241	748	738
13	**West Sumatra**	4,041	2.2	14	351	93.7	329	50	379	374	93.9	325	192	517	507
14	Riau	3,372	1.9												
15	Jambi	2,059	1.1												
16	**South Sumatra**	6,403	3.5	15	349	91.1	318	55	373	371	96.0	332	229	561	550
17	Bengkulu	1,213	0.7												
18	**Lampung**	6,108	3.4	11	274	94.5	259	38	297	297	93.8	257	164	421	414
31	**DKI Jakarta**	8,352	4.6	40	731	87.8	642	65	707	698	84.5	610	355	965	958
32	**West Java**	5,973	19.8	52	1,111	96.0	1,066	141	1,207	1,191	97.6	1,067	603	1670	1,658
33	**Central Java**	8,733	15.8	37	878	98.9	868	135	1,003	991	99.2	859	523	1382	1,362
34	**DI Yogyakarta**	2,923	1.6	22	478	94.4	451	49	500	494	92.8	438	203	641	636
35	**East Java**	32,713	18.0	45	1,044	96.2	1,004	117	1,121	1,111	99.0	1,025	462	1487	1,465
51	**Bali**	2 ,798	1.5	14	340	94.7	322	43	365	364	95.9	325	160	485	482
52	**West Nusa Tenggara**	3,416	1.9	16	407	98.8	402	54	456	447	99.5	396	278	674	668
53	East Nusa Tenggara	3,306	1.8												
54	East Timor	762	0.4												
61	West Kalimantan	3,292	1.8												
62	Central Kalimantan	1,431	0.8												
63	**South Kalimantan**	2,636	1.5	13	323	91.6	296	51	347	344	95.6	306	202	508	488
64	East Kalimantan	1,930	1.1												
71	North Sulawesi	2,504	1.4												
72	Central Sulawesi	1,735	1.0												
73	**South Sulawesi**	7,045	3.9	16	375	95.7	359	36	395	392	94.6	352	163	515	509
74	Southeast Sulawesi	1,382	0.8												
81	Maluku	1,885	1.0												
82	Irian Jaya	1,671	0.9												
	Total	181,548	100.0	321	7,224	94.4	6,820	878	7,698	7,619	95.2	6,799	3,775	10,574	10,435

		1990 Population			IFLS 1 HH Ivwd	IFLS4 Households: Interviewed, died, or joined other hh					IFLS5 Households: Interviewed, died, or joined other hh					Dynasties contacted	
	Provinces [a]	N(000)[b]	%	IFLS EAs		% IFLS1 HH[c]	# IFLS1 HH[d]	Any split-off HH	Total	Inter-viewed	% IFLS1 HH	# IFLS1 HH	Any split-off HH	Total	Interviewed	#	%
11	Aceh	3,476	1.9														
12	**North Sumatra**	10,391	5.7	26	563	87.6	493	532	1025	998	86.1	483	899	1,382	1,335	497	88.3
13	**West Sumatra**	4,041	2.2	14	351	89.5	314	421	735	714	88.7	305	542	847	791	321	91.5
14	Riau	3,372	1.9														
15	Jambi	2,059	1.1														
16	**South Sumatra**	6,403	3.5	15	349	86.2	301	435	736	712	85.5	295	620	915	882	308	88.3
17	Bengkulu	1,213	0.7														
18	**Lampung**	6,108	3.4	11	274	93.4	256	329	585	569	94.1	257	503	760	733	263	96.0
31	**DKI Jakarta**	8,352	4.6	40	731	75.4	551	637	1,188	1,147	68.5	492	748	1,240	1,170	540	73.9
32	**West Java**	35,973	19.8	52	1,111	93.4	1,038	1,227	2,265	2,207	90.7	991	1,643	2,634	2,496	1,035	93.2
33	**Central Java**	28,733	15.8	37	878	95.7	840	973	1,813	1,733	97.8	846	1,410	2,256	2,164	860	97.9
34	**DI Yogyakarta**	2,923	1.6	22	478	91.0	435	382	817	786	89.2	420	542	962	926	430	90.0
35	**East Java**	32,713	18.0	45	1,044	96.6	1,009	932	1,941	1,869	96.3	992	1,318	2,310	2,204	1,015	97.2
51	**Bali**	2,798	1.5	14	340	92.9	316	330	646	625	92.6	313	485	798	765	318	93.5
52	**West Nusa Tenggara**	3,416	1.9	16	407	98.0	399	484	883	858	99.8	400	786	1,186	1,147	406	99.8
53	East Nusa Tenggara	3,306	1.8														
54	East Timor	762	0.4														
61	West Kalimantan	3,292	1.8														
62	Central Kalimantan	1,431	0.8														
63	**South Kalimantan**	2,636	1.5	13	323	93.8	303	376	679	653	91.9	294	475	769	739	302	93.5
64	East Kalimantan	1,930	1.1														
71	North Sulawesi	2,504	1.4														
72	Central Sulawesi	1,735	1.0														
73	**South Sulawesi**	7,045	3.9	16	375	90.9	341	341	682	664	92.5	344	527	871	852	352	93.9
74	Southeast Sulawesi	1,382	0.8														
81	Maluku	1,885	1.0														
82	Irian Jaya	1,671	0.9														
	Total	1,548	100.0	321	7,224	91.3	6,596	7399	13,995	13,535	90.2	6,432	10,498	16,930	16,204	6,647	92.0

a. Boldface denotes IFLS provinces in 1993. In 1999, East Timor voted for independence from Indonesia and became the soverign state of Timor Leste. Also since 1999, a number of new provinces has been formed.
b. Source of population number is the BPS 1990 Population Census.
c. The percentage is out of IFLS1 HH with at least some members living in the last survey.
d. Includes IFLS1 HH whose members had all died or joined other IFLS households by the time of the survey.

Source: IFLS5 User Guide

Appendix II: Log Linear Relationship between Wages (w) and Education (X)

W(X, t) is defined as the wages at time t of an individual with X years of schooling. Thus, the present value of wages of an individual who enters the labour market after X years of schooling is represented by the integral of wages with respect to time (t):

$PV(X) = \int_X^T W(X,t)\, e^{-rt} dt,$

where:

PV(X): present value of an individual wages; X: years of schooling; T: retirement time, and e: exponential value (approximately 2.718).

Based on assumption 1, W(X, t) does not depend on t, and we can write just W(X) and take it out the integral, thus we have PV(X) now:

$P(X) = \int_X^T W(X)\, e^{-rt} dt,$

$PV(X) = W(X) \int_X^T e^{-rt}\, dt,$

$PV(X) = W(X) \left(-\frac{1}{r}\right) [e^{-rT} - e^{-rX}],$

$PV(X) = W(X) \frac{[e^{-rX} - e^{-rT}]}{r};$

Meanwhile, according to assumption 2 of Mincer wage equation that present value of lifetime incomes are the same across individuals regardless of schooling, PV(X) should not depend on X, thus PV(X)=PV. Assumption 3 also states that the number of years spent at work are independent of the number of years schooling, thus the number of year in workforce is equal to retirement time – years of schooling WF=T-X; or we can write as: T=X+WF. Using both assumptions, now we have:

$PV(X) = PV = W(X) \frac{[e^{-rX} - e^{-rT}]}{r},$

$rPV = W(X)[e^{-rX} - e^{-r} e^{-rWF}],$

And the number of schooling should be the same regardless of the years of schooling:

$rPV = W(0)[1 - e^{-rWF}] = W(X)[e^{-rX} - e^{-rX} e^{-rWF}],$

$rPV = W(X) e^{-rX}[1 - e^{-rWF}]$, thus:

$W(0) = W(X) e^{-rX}$, and finally,

$ln(X) = lnW(0) + rX,$

The equation above proves that there is a log-linear relationship between wages and schooling[73]. As Byron and Takahashi (1989) assert that there is a strong theoretical

[73] The linear specification could be not accurate, particularly if there is an assumption that individuals are heterogeneous in their preferences and wage opportunities. Thus, log wages may either be a convex or a

rationale for using log wages in a human capital wages regression, since investments in human capital, like other investments, are only undertaken as long as the rate of return (not the absolute return) on the investment exceeds the discount rate. Log-linearity of wages as a function of years of schooling is in fact a key empirical implication of the human capital model with identical individuals.

concave function of years of schooling. Another cause of this case is the presence of ‘‘credential’’ or ‘‘sheepskin’’ effects.

Appendix III: Experience and Wage Relationship

For post-schooling investment model or experience variable, the other assumptions is needed: (4) the return to post-schooling investment in a constant δ, and (5) an employee devotes a fraction H of time to investment in the human capital, and a fraction (1-H) to actual work. This implies that growth in wages is determined by:

$$\frac{\vartheta W(X,t)}{\vartheta t} = \delta H(t)W(X,t),$$

By solving the differential equation and we will have:

$$\frac{1}{W(X,t)}\frac{\vartheta W(X,t)}{\vartheta t} = \delta H(t),$$

$$\frac{\vartheta lnW(X,t)}{\vartheta t} = \delta H(t),$$

$\int \frac{\vartheta lnW(X,t)}{\vartheta t} = \int \delta H(t)$, and

$$lnW(X,t) = \delta \int H(t).$$

For specific solution, now we have:

$$lnW(X,t) = \gamma + \delta \int_0^t H(u)du,$$

where: γ is constant,

to find the value of γ, form the assumption that there is no post-schooling investment in human capital, so:

$lnW(X) = lnW(0) + rX$ or equation that we obtain in Appendices I,

if H(u)=0 and put into differential equation:

$lnW(X,t) = \gamma$, this implies that $\gamma = lnW(0) + rX$, and

the solution is:

$$lnW(X,t) = lnW(0) + rX + \delta \int_0^t H(u)du.$$

It starts from time 0 because the growth does not start till after school is over. To solve the model, we assumed that the frequency of investment in human capital is as follow:

$$H(t) = H^*\left(1 - \frac{t}{T}\right) = H^* - \frac{H^*t}{T},$$

this is consistent with t running from the time an individual finishes school. Afterwards, we can figure out the relationship between potential earning and schooling variable; back again to the solution:

$$lnW(X,t) = lnW(0) + rX + \delta \int_0^t H(u)du,$$

$$lnW(X,t) = lnW(0) + rX + \delta H^*\left[u - \frac{u^2}{2T}\right]_0^t,$$

and finally the relationship between potential earning and schooling variable is:

$lnW(X,t) = lnW(0) + rX + \delta H^* t - \delta H^* \frac{t^2}{2T}$.

Since an individual only works part time now, so:

$W_{part-time}(X,t) = \big(1 - H(t)\big)W(X,t)$,

$W_{part-time}(X,t) = lnW(X,t) + ln\big(1 - H(t)\big)$, and

$ln\big(1 - H(t)\big) = \ln\left(1 - H^* + H^* \frac{t}{T}\right)$,

Using Taylor expansion around t=T to solve the model:

$$\ln\left(1 - H^* + H^* \frac{t}{T}\right) = \ln(1) + \frac{H^*/T}{1}(t - T) - \frac{1}{2}(H^*/T)^2(t - T)^2$$

$$= \frac{H^*}{T}t - H^* - \frac{1}{2}(H^*/T)^2 t^2 - \frac{1}{2}(H^*/T)^2 T^2 + (H^*/T)^2 Tt,$$

$$= \frac{H^*}{T}t - H^* - \frac{1}{2}(H^*/T)^2 t^2 - \frac{1}{2}(H^*)^2 + \frac{(H^*)^2}{T}t,$$

$$= -H^* - \frac{1}{2}(H^*)^2 - \frac{1}{2}(H^*/T)^2 t^2 + t\left[\frac{H^*}{T} + \frac{(H^*)^2}{T}\right],$$

the error is highest when $-H^* + H^* \frac{t}{T}$ is big and t=0. So, in total, wage (w) would be:

$$lnw(X,t) = lnW(X,t) + \ln\,(1 - H(t))$$

$$= lnW(0) + rX + \delta H^* t - \delta H^* \frac{t^2}{2T} - H^* + H^* \frac{t}{T} - \frac{1}{2}(H^*)^2 - \frac{1}{2}(H^*/T)^2 t^2 + t\left[\frac{H^*}{T} + \frac{(H^*)^2}{T}\right],$$

$$= \left\{lnW(0) - H^* - \frac{1}{2}(H^*)^2\right\} + rX + t\left[\frac{H^*}{T} + \frac{(H^*)^2}{T} + pH^*\right] + t^2\left\{\frac{1}{2}(H^*/T)^2 - \frac{pH^*}{2T}\right\},$$

so we can prove that standard Mincer wage equation is a linear function, because of the homogeneous individuals. Meanwhile, the quadratic function of experience variable could capture the fact that on-the-job training investments decline over time in a standard lifecycle of the human capital model; thus, more consistent with the theory and much easier for the regression model to be estimated.

Appendix IV: Endogeneity Problem in Mincer Wage Equation

This part will prove the endogeneity by mathematical equation based on Card (2001). Suppose the model of the observed and wage outcome is the semi-logarithmic functional form, as adopted from Mincer Equation (1974), the model is as follow:

$\log W_{n,t} = \beta_{0,t} + \beta_{1,t}X_{n,t} - \frac{1}{2}\beta_{2,t}X_{n,t}^2 + \epsilon_{n,t},$

or it can be written as:

$\log W_{n,t} = B_{0,t} + \overline{\beta_{1,t}}X_{n,t} - \frac{1}{2}\beta_2 X_{n,t}^2 + B_{n,t} + (\beta_1 - \overline{\beta_1})X_{n,t} + \epsilon_{n,t},$

where:

$\beta_{0,t}$: a person-specific constant of integration, and $B_{n,t} \equiv \beta_{0,t} - B_{0,t}$,

$\epsilon_{n,t}$: error term, we can ignore it in this case.

If the model assumed heterogeneity of individuals, it would allow to affect the intercept of the wage equation (via $\beta_{0,t}$) and the slope of the wage-schooling relation (via $\beta_{1,t}$), and both equations are in terms of underlying random variables $\beta_{0,t}$, $\beta_{1,t}$ and MC_x (the marginal cost of schooling). In estimating the OLS, the model ignores other covariates. Considering the linear projection of $\beta_{0,t}$ and $\left(\beta_{1,t} - \underline{\beta_{1,t}}\right)$ on observed schooling:

$b_{1,t} = \lambda_0\left(X_{n,t} - \underline{X}\right) + u_{n,t}$, thus,

$\beta_{1,t} - \underline{\beta_{1,t}} = \psi_0\left(X_{n,t} - \underline{X}\right) + \xi_{n,t},$

where:

$\underline{X}$: the mean of schooling, and if expected correlation between $X_{n,t}$ and $u_{n,t}$ or $\xi_{n,t}$ (error term) is zero or:

$E\left[X_{n,t}u_{n,t}\right] = E\left[X_{n,t}\xi_{n,t}\right] = 0,$

then, substituting these expressions to the wage function:

$log\ log\ W_{\ n,t} = Constant + \left(\underline{\beta_1} + \lambda_0 - \psi_0\underline{X}\right)X_{n,t} + \left(\psi_0 - \frac{1}{2}\right)X_{n,t}^2 + u_{n,t} + \xi_n X_{n,t}.$

From the equation above, the orthogonality of $\xi_{n,t}$ and $X_{n,t}$ (the last term) does not imply that $E\left[\xi_{n,t}X_{n,t}^2\right] = 0$; as a result, the residual component $\xi_{n,t}X_{n,t}$ may be correlated with schooling ($X_{n,t}$). Yet, if the third central moments of the joint distribution of β_1 and $MC_{x(n,t)}$ are equal to zero, then $\xi_{n,t}X_{n,t}$will be uncorrelated with the education variable ($X_{n,t}$).

Furthermore, if $E\left[\left(X_{n,t}-\underline{X}\right)^{3}\right]=0$ implies that the linear projection of $X_{n,t}^{2}$ on $X_{n,t}$ has slope $2\underline{X}$. Under this assumption, the probability limit of the ordinary least square (OLS) regression coefficient b_{ols} from a regression of log wages on schooling is:

$plim\ \beta_{1,ols}=\left(\underline{\beta_1}+\lambda_0-\psi_0\underline{X}\right)+2\underline{X}\left(\psi_0-\frac{1}{2}\beta_2\right),$

$$plim\ \beta_{1,ols}=\underline{\beta_1}+\lambda_0-\psi_0\underline{X}+2\psi_0\underline{X}-\underline{X}\beta_2$$

$plim\ \beta_{1,ols}=\underline{\beta_1}-\beta_2\underline{X}+\lambda_0+\psi_0\underline{X},$

$plim\ \beta_{1,ols}=\underline{\beta_1}+\lambda_0+\psi_0\underline{X},$

the equation above generalises the conventional analysis of ability bias in the relationship between schooling and wage.

Suppose that there is no heterogeneity in the marginal benefits of schooling or $\beta_1=\underline{\beta_1}$; and the log wages are linear in schooling $\beta_2=0$, then:

$plim\ \beta_{1,ols}-\underline{\beta_1}=\lambda_0,$

this is the standard expression for the asymptotic bias in the estimated return to schooling. This arises by applying the omitted variables formula to the wage model with a constant schooling coefficient $\underline{\beta_1}$. Based on the model, the bias occurs through the correlation between the ability component $b_{n,t}$ and the marginal cost of schooling $MC_{X(n,t)}$. If the marginal costs are lower for individuals who would tend to receive more at any level of schooling, then the standard deviation is less then zero ($\sigma_x<0$), which implies that $\lambda_0>0$.

Appendix V: Education Attainment Based on Gender

Table V.1: Education Attainment Based on Gender, 2000 and 2014

Highest level of education attended	IFLS3			IFLS5		
	sex		Total	sex		Total
	Male	Female		Male	Female	
No Schooling	1	2	3			
	0.01	*0.02*	*0.01*			
Elementary school	4,461	5,041	9,502	4,528	5,368	9,896
	39.81	*44.98*	*42.39*	*28.36*	*32.12*	*30.28*
Junior high general	1,700	1,569	3,269	2,529	2,595	5,124
	15.17	*14*	*14.58*	*15.84*	*15.53*	*15.68*
Junior high vocational	196	136	332	126	80	206
	1.75	*1.21*	*1.48*	*0.79*	*0.48*	*0.63*
Senior high general	1,645	1,497	3,142	2,706	2,635	5,341
	14.68	*13.36*	*14.02*	*16.95*	*15.77*	*16.34*
Senior high vocational	1,460	1,210	2,670	2,521	2,025	4,546
	13.03	*10.8*	*11.91*	*15.79*	*12.12*	*13.91*
Adult education A	4	12	16	12	10	22
	0.04	*0.11*	*0.07*	*0.08*	*0.06*	*0.07*
Adult education B	4	2	6	59	49	108
	0.04	*0.02*	*0.03*	*0.37*	*0.29*	*0.33*
Adult education C				154	86	240
				0.96	*0.51*	*0.73*
Open university	5	4	9	10	18	28
	0.04	*0.04*	*0.04*	*0.06*	*0.11*	*0.09*
Islamic school	41	32	73	41	37	78
	0.37	*0.29*	*0.33*	*0.26*	*0.22*	*0.24*
School for disable	1	1	2	7	6	13
	0.01	*0.01*	*0.01*	*0.04*	*0.04*	*0.04*
Diploma	411	432	843	521	705	1,226
	3.67	*3.85*	*3.76*	*3.26*	*4.22*	*3.75*
Undergraduate	694	460	1,154	1,597	1,595	3,192
	6.19	*4.1*	*5.15*	*10*	*9.54*	*9.77*
Master	26	10	36	114	88	202
	0.23	*0.09*	*0.16*	*0.71*	*0.53*	*0.62*
PhD				8	5	13
				0.05	*0.03*	*0.04*
Islamic Elementary School (Madrasah	162	294	456	160	243	403
	1.45	*2.62*	*2.03*	*1*	*1.45*	*1.23*
Islamic Junior/High School (Madrasah	274	348	622	457	643	1,100
	2.44	*3.1*	*2.77*	*2.86*	*3.85*	*3.37*
Islamic Senior/High School (Madrasah	115	152	267	406	503	909
	1.03	*1.36*	*1.19*	*2.54*	*3.01*	*2.78*
Kindergarten	0	2	2	0	1	1
	0	*0.02*	*0.01*	*0*	*0.01*	*0*
Other	7	4	11	5	11	16
	0.06	*0.04*	*0.05*	*0.03*	*0.07*	*0.05*
Don't know				5	6	11
				0.03	*0.04*	*0.03*
Missing				0	2	2
				0	*0.01*	*0.01*
Total	**11,207**	**11,208**	**22,415**	**15,966**	**16,711**	**32,677**
	100	***100***	***100***	***100***	***100***	***100***

Source: The author's calculation.

Appendix VI: Mincer Wage Equation, OLS

Table VI. 1: Estimations: All Individuals (2000)

	All Individuals					
2000	**Basic + Personal**			**Basic + personal + work related + firm size**		
Log of real hourly wage	**Coef.**	**Std. Err.**	**P>t**	**Coef.**	**Std. Err.**	**P>t**
Junior High School	0.365	0.034	***	0.293	0.034	***
Senior High School	0.825	0.030	***	0.659	0.033	***
University	1.549	0.038	***	1.210	0.044	***
Experience	0.048	0.005	***	0.031	0.005	***
Experience squared	-0.001	0.000	***	-0.001	0.000	***
Sex (1=female)	-0.304	0.025	***	-0.351	0.025	***
Married and cohabitate	0.105	0.035	***	0.087	0.034	***
Other (Separated, divorced and widowed)	-0.089	0.068		-0.064	0.065	
Religion1: Islam	0.003	0.062		0.012	0.060	
Religion2: Christian/Protestant	0.066	0.090		0.089	0.086	
Religion3: Catholic	0.031	0.104		0.046	0.099	
Religion5: Buddhist	-0.307	0.186	*	-0.168	0.178	
Ethnicity1: Jawa	-0.002	0.032		0.021	0.031	
Ethnicity2: Sunda	0.172	0.040	***	0.175	0.039	***
Ethnicity3: Batak	0.282	0.081	***	0.314	0.078	***
Ethnicity4: Betawi	0.311	0.051	***	0.343	0.050	***
Ethnicity5: Minang	0.200	0.056	***	0.208	0.054	***
Ethnicity6: Tiong Hoa	0.266	0.162		0.328	0.156	**
Status: fulltime (30 hours a week or more)				-0.621	0.032	***
Tenure				0.019	0.005	***
Tenure squared				-0.000	0.000	
Sector: private				-0.268	0.038	***
Industry2: mining and quarrying				0.406	0.121	***
Industry3: manufacturing				0.140	0.039	***
Industry4: electricity, gas and water				0.116	0.144	
Industry5: construction				0.309	0.048	***
Industry6: wholesale, retail, restaurants and hotels				0.096	0.045	**
Industry7: transportation, storage, and communications				0.179	0.057	***
Industry8: Finance, insurance, real estate and business services				0.450	0.093	***
Industry9: Social services				0.127	0.037	***
Firm size2: 5-19 people				0.179	0.027	***
Firm size3: 20-99 people				0.254	0.032	***
Firm size4: >= 100 people				0.400	0.043	***
Constants	6.484	0.069	0.000	7.160	0.087	***
Observation		6386			6386	
The R-squared statistic		0.290			0.354	

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table VI.2: Estimations: All Individuals (2014)

2014	All Individuals					
	Basic + Personal			Basic + personal + work related + firm size		
Log of real hourly wage	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>t
Junior High School	0.298	0.039	***	0.241	0.038	***
Senior High School	0.694	0.035	***	0.525	0.035	***
University	1.251	0.038	***	0.972	0.043	***
Experience	0.037	0.004	***	0.025	0.004	***
Experience squared	-0.000	0.000	***	-0.000	0.000	***
Sex (1=female)	-0.299	0.023	***	-0.281	0.023	***
Married and cohabitate	0.054	0.035		0.012	0.033	
Other (Separated, divorced and widowed)	0.039	0.064		0.084	0.061	
Religion1: Islam	-0.144	0.052	***	-0.187	0.049	***
Religion2: Christian/Protestant	-0.011	0.107		-0.028	0.102	
Religion3: Catholic	0.002	0.089		-0.046	0.084	
Religion5: Buddhist	-0.059	0.314		-0.107	0.299	
Ethnicity1: Jawa	0.046	0.028		0.052	0.027	*
Ethnicity2: Sunda	0.193	0.038	***	0.169	0.036	***
Ethnicity3: Batak	0.143	0.069	**	0.132	0.065	**
Ethnicity4: Betawi	0.201	0.166		0.156	0.158	
Ethnicity5: Minang	0.239	0.044	***	0.263	0.042	***
Ethnicity6: Tiong Hoa	0.406	0.176	**	0.531	0.168	***
Status: fulltime (30 hours a week or more)				-0.423	0.030	***
Tenure				0.019	0.005	***
Tenure squared				0.000	0.000	
Sector: private				-0.247	0.035	***
Industry2: mining and quarrying				0.412	0.088	***
Industry3: manufacturing				0.055	0.045	
Industry4: electricity, gas and water				0.200	0.114	*
Industry5: construction				0.266	0.063	***
Industry6: wholesale, retail, restaurants and hotels				0.006	0.046	
Industry7: transportation, storage, and communications				0.118	0.073	
Industry8: Finance, insurance, real estate and business services				0.305	0.057	***
Industry9: Social services				-0.136	0.044	***
Firm size2: 5-19 people				0.065	0.030	**
Firm size3: 20-99 people				0.270	0.032	***
Firm size4: >= 100 people				0.578	0.035	***
Constants	7.884	0.064	***	8.962	0.132	***
Observation	8119			8119		
The R-squared statistic	0.178			0.260		

Source: The author’s calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table VI.3: Estimation Results: By Gender (2000)

2000	By Gender											
	Male			Female			Male			Female		
	Coef.	SE	P>t	Coef.	SE	P>t	Coef.	SE	P>t	Coef.	SE	P>t
Junior High School	0.336	0.040	***	0.393	0.066	***	0.291	0.040	***	0.243	0.064	***
Senior High School	0.710	0.036	***	1.063	0.056	***	0.593	0.039	***	0.790	0.060	***
University	1.399	0.048	***	1.796	0.063	***	1.130	0.054	***	1.312	0.077	***
Experience	0.052	0.006	***	0.044	0.008	***	0.042	0.006	***	0.018	0.008	**
Experience squared	-0.001	0.000	***	-0.001	0.000	***	-0.001	0.000	***	0.000	0.000	
Sex (1=female)	(omitted)			(omitted)			(omitted)			(omitted)		
Married and cohabitate	0.086	0.043	*	0.150	0.059	***	0.117	0.042	***	0.055	0.056	
Other (Separated, divorced and widowed)	-0.003	0.126		-0.004	0.089		0.033	0.121		-0.011	0.084	
Religion1: Islam	0.042	0.074		-0.085	0.112		0.056	0.072		-0.088	0.106	
Religion2: Christian/Protestant	0.038	0.111		0.057	0.154		0.048	0.106		0.086	0.145	
Religion3: Catholic	0.025	0.129		-0.046	0.176		0.042	0.124		-0.046	0.166	
Religion5: Buddhist	-0.221	0.220		-0.428	0.342		-0.069	0.211		-0.330	0.321	
Ethnicity1: Jawa	0.020	0.038		-0.034	0.058		0.045	0.037		-0.014	0.056	
Ethnicity2: Sunda	0.136	0.047	***	0.223	0.073	***	0.131	0.046	***	0.242	0.070	***
Ethnicity3: Batak	0.287	0.097	***	0.253	0.143	*	0.349	0.094	***	0.258	0.135	*
Ethnicity4: Betawi	0.214	0.062	***	0.471	0.090	***	0.228	0.061	***	0.518	0.086	***
Ethnicity5: Minang	0.242	0.068	***	0.082	0.099		0.237	0.066	***	0.097	0.094	
Ethnicity6: Tiong Hoa	0.318	0.199		0.178	0.280		0.375	0.192	**	0.292	0.264	
Status: fulltime (30 hours a week or more)							-0.717	0.042	***	-0.551	0.050	***
Tenure							0.011	0.006	**	0.025	0.008	***
Tenure squared							0.000	0.000		0.000	0.000	
Sector: private							-0.198	0.045	***	-0.399	0.068	***
Industry2: mining and quarrying							0.372	0.124	***	1.127	0.510	**
Industry3: manufacturing							0.170	0.048	***	0.141	0.068	**
Industry4: electricity, gas and water							0.127	0.155		0.160	0.365	
Industry5: construction							0.297	0.052	***	0.045	0.167	
Industry6: wholesale, retail, restaurants and hotels							0.100	0.056	*	0.129	0.075	*
Industry7: transportation, storage, and communications							0.184	0.060	***	-0.213	0.224	
Industry8: Finance, insurance, real estate and business services							0.396	0.115	***	0.547	0.158	***
Industry9: Social services							0.168	0.045	***	0.137	0.067	**
Firm size2: 5-19 people							0.108	0.033	***	0.321	0.050	***

Firm size3: 20-99 people							0.154	0.039	***	0.440	0.056	***
Firm size4: >= 100 people							0.282	0.054	***	0.643	0.070	***
Constants	6.484	0.083	***	6.153	0.118	***	7.130	0.102	***	6.880	0.158	***
Observation		4216			2170			4216			2170	
The R-squared statistic		0.231			0.361			0.297			0.443	

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table VI.4: Estimation Results: By Gender (2014)

2014	By Gender											
	Basic + Personal			Basic + Personal			Basic + personal + work related + firm size			Basic + personal + work related + firm size		
	Male			Female			Male			Female		
	Coef.	SE	P>t	Coef.	SE	P>t	Coef.	SE	P>t	Coef.	SE	P>t
Junior High School	0.259	0.048	***	0.359	0.068	***	0.197	0.046	***	0.308	0.065	***
Senior High School	0.635	0.042	***	0.784	0.062	***	0.464	0.043	***	0.622	0.062	***
University	1.181	0.048	***	1.336	0.063	***	0.913	0.053	***	1.02	0.075	***
Experience	0.033	0.006	***	0.044	0.007	***	0.027	0.005	***	0.026	0.007	***
Experience squared	0	0	***	-0.001	0	***	0	0	***	0	0	**
Sex (1=female)		(omitted)			(omitted)			(omitted)			(omitted)	
Married and cohabitate	0.147	0.043	***	-0.102	0.059	*	0.098	0.041	**	-0.114	0.056	**
Other (Separated, divorced and widowed)	0.164	0.098	*	-0.114	0.091		0.166	0.094	*	-0.024	0.086	
Religion1: Islam	-0.079	0.063		-0.261	0.088	***	-0.118	0.06	**	-0.289	0.084	***
Religion2: Christian/Protestant	0.014	0.136		-0.05	0.173		0.001	0.13		-0.045	0.164	
Religion3: Catholic	-0.024	0.111		0.014	0.148		-0.074	0.106		-0.017	0.14	
Religion5: Buddhist	-0.003	0.371		-0.088	0.577		0.09	0.352		-0.395	0.546	
Ethnicity1: Jawa	0.043	0.035		0.06	0.049		0.05	0.033		0.06	0.047	
Ethnicity2: Sunda	0.143	0.045	***	0.291	0.067	***	0.133	0.043	***	0.232	0.064	***
Ethnicity3: Batak	0.146	0.085	*	0.134	0.115		0.136	0.081	*	0.111	0.109	
Ethnicity4: Betawi	0.26	0.22		0.124	0.254		0.282	0.209		-0.042	0.242	
Ethnicity5: Minang	0.177	0.055	***	0.336	0.075	***	0.191	0.052	***	0.362	0.071	***
Ethnicity6: Tiong Hoa	0.344	0.209	***	0.518	0.319		0.437	0.199	**	0.652	0.304	**
Status: fulltime (30 hours a week or more)							-0.595	0.041	***	-0.294	0.044	***
Tenure							0.016	0.006	***	0.021	0.008	***
Tenure squared							0	0	***	0	0	
Sector: private							-0.236	0.044	***	-0.244	0.057	***
Industry2: mining and quarrying							0.472	0.089	***	0.584	0.38	
Industry3: manufacturing							0.156	0.052	***	-0.131	0.086	
Industry4: electricity, gas and water							0.298	0.115	***	-0.24	0.502	
Industry5: construction							0.266	0.066	***	0.623	0.232	***
Industry6: wholesale, retail, restaurants and hotels							0.106	0.054	**	-0.167	0.087	**

Industry7: transportation, storage, and communications							0.158	0.076	**	0.299	0.238	
Industry8: Finance, insurance, real estate and business services							0.358	0.064	***	0.216	0.115	*
Industry9: Social services							-0.076	0.051		-0.219	0.083	***
Firm size2: 5-19 people							0.087	0.037	**	0.012	0.05	
Firm size3: 20-99 people							0.223	0.039	***	0.333	0.055	***
Firm size4: >= 100 people							0.519	0.043	***	0.673	0.062	***
Constants	7.852	0.079	***	7.655	0.105	***	8.974	0.165	***	8.749	0.221	***
Observation	4946			3173			4946			3173		
The R-squared statistic	0.157			0.183			0.245			0.276		

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table VI.5: Estimation Results with Interaction Variables between Gender and Years of Schooling

	2000			2014		
Log of real hourly wage	**Coef.**	**Std. Err.**	**P>t**	**Coef.**	**Std. Err.**	**P>t**
Years of Schooling	0.103	0.005	***	0.090	0.005	***
Experience	0.036	0.005	***	0.022	0.004	***
Experience squared	-0.001	0.000	***	0.000	0.000	***
Sex (1=female)	-0.907	0.073	***	-0.518	0.075	***
Married and cohabitate	0.107	0.033	***	0.058	0.033	*
Other (Separated, divorced and widowed)	0.008	0.065		0.123	0.059	**
.						
.						
Urban	-0.004	0.026		0.140	0.024	***
Sex* Years of schooling	0.058	0.007	***	0.022	0.006	***
Constants	6.476	0.141	***	8.236	0.162	***
Observation	6,386			8,119		
R-squared	0.376			0.301		

Source: The author’s calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table VI.5: Estimation Results: By Sector 2000

	By Sector											
	Public			Private			Public			Private		
2000	**Coef.**	**SE**	**P>t**	**Coef.**	**SE**	**P>t**	**Coef.**	**SE**	**P>t**	**Coef.**	**SE**	**P>t**
Junior High School	0.417	0.126	***	0.295	0.037	***	0.331	0.127	***	0.270	0.036	***
Senior High School	0.950	0.099	***	0.683	0.035	***	0.825	0.105	***	0.607	0.035	***
University	1.496	0.103	***	1.366	0.052	***	1.285	0.111	***	1.207	0.053	***
Experience	0.032	0.012	***	0.038	0.005	***	0.005	0.013		0.031	0.005	***
Experience squared	0.000	0.000		-0.001	0.000	***	0.000	0.000		-0.001	0.000	***
Sex (1=female)	-0.017	0.054		-0.367	0.028	***	-0.058	0.053		-0.414	0.028	***
Married and cohabitate	0.324	0.097	***	0.112	0.037	***	0.239	0.095	***	0.083	0.036	**
Other (Separated, divorced and widowed)	0.138	0.175		-0.055	0.073		0.180	0.170		-0.074	0.070	
Religion1: Islam	-0.072	0.103		0.040	0.074		-0.014	0.100		0.000	0.071	
Religion2: Christian/Protestant	-0.201	0.158		0.179	0.105	*	-0.128	0.153		0.139	0.101	
Religion3: Catholic	-0.245	0.181		0.135	0.122		-0.296	0.175	*	0.116	0.117	
Religion5: Buddhist	0.443	0.800		-0.244	0.193		-0.007	0.773		-0.159	0.185	
Ethnicity1: Jawa	0.035	0.062		0.038	0.036		0.033	0.061		0.034	0.035	
Ethnicity2: Sunda	-0.046	0.082		0.247	0.045	***	-0.050	0.079		0.221	0.043	***
Ethnicity3: Batak	0.188	0.147		0.345	0.094	***	0.178	0.143		0.371	0.090	***
Ethnicity4: Betawi	-0.125	0.150		0.423	0.056	***	-0.117	0.145		0.393	0.054	***
Ethnicity5: Minang	0.296	0.104	***	0.187	0.065	***	0.204	0.101	**	0.191	0.062	***
Ethnicity6: Tiong Hoa	0.000	(omitted)		0.346	0.166	**	0.000	(omitted)		0.334	0.160	**
Status: fulltime (30 hours a week or more)							-0.459	0.067	***	-0.665	0.036	***
Tenure							0.043	0.011	***	0.017	0.005	***
Tenure squared							-0.001	0.000	**	0.000	0.000	
												(omitted)
Sector: private							(omitted)					
Industry2: mining and quarrying							0.853	0.547		0.370	0.125	***
Industry3: manufacturing							0.062	0.148		0.133	0.041	***
Industry4: electricity, gas and water							0.006	0.220		0.144	0.185	
Industry5: construction							0.149	0.185		0.312	0.051	***
Industry6: wholesale, retail, restaurants and hotels							-0.171	0.205		0.102	0.047	**
Industry7: transportation, storage, and communications							0.220	0.154		0.173	0.061	***

Industry8: Finance, insurance, real estate and business services							0.286	0.200		0.503	0.105	***
Industry9: Social services							0.135	0.100		0.101	0.041	**
Firm size2: 5-19 people							0.172	0.071	**	0.166	0.030	***
Firm size3: 20-99 people							0.207	0.076	***	0.268	0.036	***
Firm size4: >= 100 people							0.151	0.104		0.467	0.047	***
Constant	6.47	0.16	***	6.58	0.08	***	6.77	0.18	***	7.00	0.09	***
Observation	1060			5326			1060			5326		
The R-squared statistic	0.28			0.21			0.34			0.28		

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table VI.6: Estimation Results: By Sector 2014

2014	Basic + Personal Public Coef.	SE	P>t	Basic + Personal Private Coef.	SE	P>t	Basic + personal + work related + firm size Public Coef.	SE	P>t	Basic + personal + work related + firm size Private Coef.	SE	P>t
Junior High School	0.111	0.275		0.175	0.043	***	-0.002	0.272		0.154	0.041	***
Senior High School	0.927	0.234	***	0.538	0.039	***	0.681	0.238	***	0.439	0.039	***
University	1.542	0.233	***	0.971	0.047	***	1.271	0.241	***	0.821	0.049	***
Experience	0.056	0.014	***	0.023	0.005	***	0.066	0.014	***	0.009	0.005	*
Experience squared	0.000	0.000		0.000	0.000	***	-0.001	0.000	*	0.000	0.000	
Sex (1=female)	-0.205	0.064	***	-0.321	0.026	***	-0.178	0.063	***	-0.290	0.026	***
Married and cohabitate	0.244	0.113	**	0.094	0.039	**	0.180	0.112		0.041	0.037	
Other (Separated, divorced and widowed)	0.447	0.208	**	0.072	0.072		0.388	0.203	*	0.100	0.068	
Religion1: Islam	-0.293	0.121	**	-0.177	0.063	***	-0.282	0.118	**	-0.220	0.060	***
Religion2: Christian/Protestant	-0.428	0.248	*	0.091	0.123		-0.388	0.241		0.058	0.117	
Religion3: Catholic	-0.326	0.203		0.061	0.101		-0.358	0.198	*	0.014	0.095	
Religion5: Buddhist	1.341	1.045		-0.193	0.373		1.415	1.011		-0.304	0.353	
Ethnicity1: Jawa	0.322	0.077	***	0.079	0.036	**	0.351	0.076	***	0.040	0.035	
Ethnicity2: Sunda	0.281	0.109	***	0.269	0.046	***	0.304	0.106	***	0.196	0.044	***
Ethnicity3: Batak	0.404	0.160	**	0.109	0.080		0.412	0.157	***	0.092	0.075	
Ethnicity4: Betawi	0.272	0.295		0.463	0.062	***	0.212	0.287		0.385	0.059	***

Ethnicity5: Minang	0.486	0.099	***	0.224	0.057	***	0.575	0.098	***	0.206	0.054	***
Ethnicity6: Tiong Hoa	0.000	(omitted)		0.432	0.186	**	0.000	(omitted)		0.496	0.176	***
Status: fulltime (30 hours a week or more)							-0.302	0.069	***	-0.504	0.036	***
Tenure							-0.007	0.013		0.039	0.006	***
Tenure squared							0.001	0.000	*	-0.001	0.000	***
Sector: private							0.000	(omitted)		0.000	(omitted)	
Industry2: mining and quarrying							0.490	0.284	*	0.339	0.099	***
Industry3: manufacturing							-0.402	0.289		0.019	0.049	
Industry4: electricity, gas and water							-0.356	0.338		0.206	0.131	
Industry5: construction							-0.116	0.598		0.259	0.067	***
Industry6: wholesale, retail, restaurants and hotels							-0.449	0.440		-0.054	0.049	
Industry7: transportation, storage, and communications							0.656	0.410		0.049	0.079	
Industry8: Finance, insurance, real estate and business services							0.326	0.216		0.242	0.062	***
Industry9: Social services							-0.203	0.149		-0.196	0.049	***
Firm size2: 5-19 people							0.180	0.113		0.063	0.033	*
Firm size3: 20-99 people							0.403	0.113	***	0.262	0.036	***
Firm size4: >= 100 people							0.724	0.126	***	0.584	0.040	***
Constants	7.102	0.269	***	8.136	0.074	***	7.368	0.302	***	8.539	0.084	***
Observation	1220			5942			1220			5942		
The R-squared statistic	0.251			0.131			0.309			0.225		

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Appendix VII: Robustness Test

Table VII.1: Natural and Conventional Instruments

Variable	2000			2014		
	Number of Obs	**Question**	**Book**	**Number of Obs**	**Question**	**Book**
Total observation	***6386***			***8119***		
	Endogeneity problem					
Conventional Instruments						
Total: National exam score	933	DL16e	BIIIA	630	DL16e	BIIIA
Siblings or twin data	3888	BA28	BIIIB	7021		
Do you have siblings out of hh? (YES?NO)	2043	BA29a	BIIIB	6331		
# of siblings away who died lst/12 months	4538	BA30	BIIIB			
Total # of siblings non-HHM						
Siblings' education	4251	BA36	BIIIB			
Distance to school(in hour and minutes)	2031	DL16j	BIIIA	2267	DL16j	BIIIA
	2031	dl16jx				
Spouse education	3062	KW02n	BIIIA	2	KW02n	BIIIA
Parent's highest education						
Mother's education	2857	BA08M	BIIIB	4063	BA08M	BIIIB
Father's education	3769	BA08P	BIIIB	4842	BA08P	BIIIB
Preschool attendance				7067	dl05b	BIIIA
Delayed enrolment in primary school (age of primary school enrolment)	6371	dl05a	BIIIA	7058	dl05a	BIIIA
In what age did you enter school (DL05A)						
Smoking	6383	KM01a	BIIIB	7036	KM01a	BIIIB
Policy instruments (Purnastuti et al (2015)						
INPRES Program (Dummy for INPRES program)				1 if individual was born in 1967 and later; 0 otherwise		
The first compulsory school attendance law (CSAL-1),				1 if individual was born in 1977 and later		
The second compulsory school attendance law (CSAL-2).				1 if individual was born in 1987 and later		
	Selection problem					
Heckman sample selection model						
Number of children		AR02b	Book K			
Dumauli (2015)	Household size					

Source: The author's compilation.

Appendix VIII: IV Model of Mincer Wage Equation

Table VIII.1: Age Distribution and Year of Schooling (in per cent)

Age in 2000	Year of Schooling 6	9	12	15	Total	Age in 2014	Year of Schooling 6	9	12	15	Total
16	2.17	3.70	0.00	0.00	1.46	16	1.00	1.10	0.00	0.00	0.36
17	2.74	7.23	0.00	0.00	2.30	17	1.21	2.57	0.00	0.00	0.64
18	2.61	7.06	1.92	0.00	2.82	18	0.85	2.28	2.34	0.00	1.45
19	2.74	5.59	4.15	0.00	3.29	19	0.93	3.31	4.33	0.00	2.43
20	3.91	4.04	4.90	0.00	3.66	20	1.21	2.57	3.58	0.00	2.06
21	3.13	4.13	5.16	0.00	3.48	21	1.14	2.13	3.93	0.00	2.11
22	4.13	4.99	4.85	2.75	4.31	22	1.71	3.68	4.27	2.89	3.36
23	3.04	4.48	4.04	4.33	3.81	23	1.78	2.43	3.77	3.08	3.02
24	3.09	3.44	4.50	2.96	3.57	24	1.92	3.16	3.77	3.82	3.36
25	3.87	3.27	4.30	3.80	3.88	25	2.21	3.75	3.93	4.76	3.82
26	2.57	3.87	3.84	3.70	3.37	26	2.14	2.87	3.49	4.01	3.29
27	2.87	5.77	3.54	4.65	3.87	27	1.78	3.38	4.30	4.52	3.77
28	3.13	2.84	4.90	3.59	3.70	28	2.21	2.72	3.77	4.66	3.56
29	2.44	2.93	4.70	4.22	3.49	29	2.92	5.00	3.89	4.34	4.03
30	3.13	3.10	4.60	4.12	3.73	30	3.06	4.04	4.08	5.08	4.16
31	2.44	1.98	3.94	5.70	3.30	31	2.92	6.18	3.86	4.06	4.14
32	2.74	2.84	3.99	4.22	3.37	32	3.28	4.93	4.36	4.71	4.36
33	2.57	1.89	3.64	3.80	2.96	33	3.77	3.24	3.86	4.99	4.04
34	2.17	1.98	3.44	4.33	2.85	34	3.85	4.19	3.49	3.78	3.74
35	3.13	1.81	2.98	4.01	2.98	35	3.21	4.19	2.74	3.92	3.37
36	2.26	2.41	2.53	3.70	2.58	36	3.92	3.01	3.05	3.68	3.36
37	3.61	1.98	1.82	3.70	2.77	37	3.85	3.38	2.65	3.26	3.14
38	2.91	1.12	2.38	3.70	2.54	38	3.92	2.57	2.46	2.38	2.71
39	3.04	1.20	2.02	2.96	2.38	39	4.84	2.13	2.55	2.80	2.94
40	2.87	1.46	1.97	3.48	2.43	40	3.06	3.01	2.49	1.59	2.44
41	2.83	1.64	2.12	2.85	2.40	41	2.99	2.35	1.96	2.28	2.29
42	3.52	1.29	1.97	3.27	2.60	42	2.92	2.21	2.52	1.63	2.30
43	2.22	1.29	1.11	3.27	1.86	43	2.35	1.69	1.56	2.24	1.90
44	2.17	1.20	0.76	2.01	1.53	44	2.49	1.47	2.46	2.38	2.28
45	1.87	0.95	1.01	2.53	1.53	45	2.85	1.47	1.96	2.28	2.12
46	1.78	1.20	1.06	1.90	1.47	46	2.78	1.40	1.15	2.00	1.70
47	2.22	1.29	1.26	2.32	1.77	47	2.85	1.10	1.25	2.10	1.72
48	1.70	1.12	1.26	1.37	1.41	48	1.42	0.88	1.09	2.10	1.38
49	1.91	1.03	0.71	0.74	1.21	49	2.99	1.10	1.31	1.63	1.65
50	1.30	1.20	0.71	1.90	1.19	50	2.21	0.96	0.93	1.77	1.38
51	0.83	0.60	1.01	0.95	0.86	51	2.78	0.44	0.90	1.54	1.32
52	1.44	0.69	0.76	0.53	0.96	52	2.14	0.96	0.56	1.77	1.22
53	1.17	0.69	0.81	0.95	0.94	53	2.14	0.59	0.50	1.35	1.02

54	0.52	0.52	0.91	0.95	0.70
55	1.17	0.17	0.40	0.74	0.69
Total	100.00	100.00	100.00	100.00	100.00

54	2.49	0.88	0.44	1.31	1.10
55	1.92	0.66	0.44	1.31	0.96
Total	100.00	100.00	100.00	100.00	100.00

Source: The author's calculation.

Note: The birth cohorts are in the grey area, or the age range that covered in the present study were from 16-41 years old in 2000 and from 30-55 years old in 2014. While, workers aged 42-55 were too old or entered pension age in the later period, and aged 16-29 were too young in 2014, they have not yet entered labour market in 2000.

Table VIII.2: Summary Statistics of IV Instruments

Education Level	**2000: INPRES**			**2014: CSAL1**			**2014: CSAL2**		
Percentage of distribution	**0**	**1**	**Total**	**0**	**1**	**Total**	**0**	**1**	**Total**
Primary School	**1,073**	**1,226**	**2,299**	**690**	**714**	**1,404**	**1,153**	**251**	**1,404**
	42.4%	*31.8%*	*36.0%*	*26.2%*	*13.0%*	*17.3%*	*20.2%*	*10.4%*	*17.3%*
Junior High School	**312**	**850**	**1,162**	**352**	**1,008**	**1,360**	**908**	**452**	**1,360**
	12.3%	*22.1%*	*18.2%*	*13.4%*	*18.4%*	*16.8%*	*15.9%*	*18.8%*	*16.8%*
Senior High School	**653**	**1,325**	**1,978**	**852**	**2,358**	**3,210**	**2,000**	**1,210**	**3,210**
	25.8%	*34.4%*	*31.0%*	*32.4%*	*43.0%*	*39.5%*	*35.0%*	*50.3%*	*39.5%*
University	**494**	**453**	**947**	**739**	**1,406**	**2,145**	**1,650**	**495**	**2,145**
	19.5%	*11.8%*	*14.8%*	*28.1%*	*25.6%*	*26.4%*	*28.9%*	*20.6%*	*26.4%*
Total	**2,532**	**3,854**	**6,386**	**2,633**	**5,486**	**8,119**	**5,711**	**2,408**	**8,119**
	100.0%	*100.0%*	*100.0%*	*100.0%*	*100.0%*	*100.0%*	*100.0%*	*100.0%*	*100.0%*

Source: The author’s calculation.

Table VIII.3: Age Distribution of IV Instruments

	2000								2014								
	INPRES			CSAL1			CSAL2		INPRES			CSAL1			CSAL2		
Age	0	1	Total	0	1	Total	0	Total	0	1	Total	0	1	Total	0	1	Total
16	0	93	93	0	93	93	93	93	0	29	29	0	29	29	0	29	29
17	0	147	147	0	147	147	147	147	0	52	52	0	52	52	0	52	52
18	0	180	180	0	180	180	180	180	0	118	118	0	118	118	0	118	118
19	0	210	210	0	210	210	210	210	0	197	197	0	197	197	0	197	197
20	0	234	234	0	234	234	234	234	0	167	167	0	167	167	0	167	167
21	0	222	222	0	222	222	222	222	0	171	171	0	171	171	0	171	171
22	0	275	275	0	275	275	275	275	0	273	273	0	273	273	0	273	273
23	0	243	243	0	243	243	243	243	0	245	245	0	245	245	0	245	245
24	0	228	228	228	0	228	228	228	0	273	273	0	273	273	0	273	273
25	0	248	248	248	0	248	248	248	0	310	310	0	310	310	0	310	310
26	0	215	215	215	0	215	215	215	0	267	267	0	267	267	0	267	267
27	0	247	247	247	0	247	247	247	0	306	306	0	306	306	0	306	306
28	0	236	236	236	0	236	236	236	0	289	289	0	289	289	289	0	289
29	0	223	223	223	0	223	223	223	0	327	327	0	327	327	327	0	327
30	0	238	238	238	0	238	238	238	0	338	338	0	338	338	338	0	338
31	0	211	211	211	0	211	211	211	0	336	336	0	336	336	336	0	336
32	0	215	215	215	0	215	215	215	0	354	354	0	354	354	354	0	354
33	0	189	189	189	0	189	189	189	0	328	328	0	328	328	328	0	328
34	182	0	182	182	0	182	182	182	0	304	304	0	304	304	304	0	304
35	190	0	190	190	0	190	190	190	0	274	274	0	274	274	274	0	274
36	165	0	165	165	0	165	165	165	0	273	273	0	273	273	273	0	273
37	177	0	177	177	0	177	177	177	0	255	255	0	255	255	255	0	255
38	162	0	162	162	0	162	162	162	0	220	220	220	0	220	220	0	220
39	151	0	151	151	0	151	151	151	0	239	239	239	0	239	239	0	239
40	155	0	155	155	0	155	155	155	0	198	198	198	0	198	198	0	198
41	153	0	153	153	0	153	153	153	0	186	186	186	0	186	186	0	186
42	166	0	166	166	0	166	166	166	0	187	187	187	0	187	187	0	187
43	119	0	119	119	0	119	119	119	0	154	154	154	0	154	154	0	154

44	98	0	98	98	0	98	98	98	0	185	185	185	0	185	185	0	185
45	98	0	98	98	0	98	98	98	0	172	172	172	0	172	172	0	172
46	94	0	94	94	0	94	94	94	0	138	138	138	0	138	138	0	138
47	113	0	113	113	0	113	113	113	0	140	140	140	0	140	140	0	140
48	90	0	90	90	0	90	90	90	112	0	112	112	0	112	112	0	112
49	77	0	77	77	0	77	77	77	134	0	134	134	0	134	134	0	134
50	76	0	76	76	0	76	76	76	112	0	112	112	0	112	112	0	112
51	55	0	55	55	0	55	55	55	107	0	107	107	0	107	107	0	107
52	61	0	61	61	0	61	61	61	99	0	99	99	0	99	99	0	99
53	60	0	60	60	0	60	60	60	83	0	83	83	0	83	83	0	83
54	45	0	45	45	0	45	45	45	89	0	89	89	0	89	89	0	89
55	44	0	44	44	0	44	44	44	78	0	78	78	0	78	78	0	78
Total	2,531	3,854	6,385	4,781	1,604	6,385	6,385	6,385	814	7,305	8,119	2,633	5,486	8,119	5,711	2,408	8,119

Source: The author's calculation

Table VIII.4: OLS and IV Specifications without Experience, All Individuals

	OLS			IV: First Stage			IV: Second Stage		
	Coef.	SE	P>t	Coef.	SE	P>t	Coef.	SE	P>z
Years of schooling	0.111	0.004	***				-0.004	0.038	
Sex (1=female)	-0.344	0.025	***	-0.269	0.074	***	-0.374	0.028	***
Married and cohabitate	0.245	0.027	***	-0.301	0.087	***	0.180	0.036	***
Other (Separated, divorced and widowed)	0.114	0.060	*	-1.381	0.186	***	-0.085	0.091	
Religion1: Islam	-0.009	0.107		-0.463	0.325		-0.066	0.114	
Religion2: Christian/Protestant	0.112	0.122		0.782	0.369	**	0.192	0.131	
Religion3: Catholic	0.075	0.128		1.020	0.390	***	0.182	0.140	
Religion5: Buddhist	-0.144	0.199		-0.522	0.606		-0.219	0.212	
Ethnicity1: Jawa	-0.045	0.047		0.056	0.144		-0.041	0.050	
Ethnicity2: Sunda	-0.060	0.058		-0.340	0.175	*	-0.101	0.062	
Ethnicity3: Batak	0.062	0.094		0.620	0.287	**	0.133	0.102	
Ethnicity4: Betawi	0.005	0.067		0.241	0.203		0.032	0.071	
Ethnicity5: Minang	0.048	0.094		1.106	0.285	***	0.174	0.107	
Ethnicity6: Tiong Hoa	0.114	0.158		0.346	0.480		0.146	0.167	
Status: fulltime (30 hours a week or more)	-0.646	0.032	***	-0.087	0.097		-0.654	0.034	***
Tenure	0.032	0.004	***	0.037	0.014	***	0.034	0.005	***
Tenure squared	-0.001	0.000	***	-0.002	0.000	***	-0.001	0.000	***
Sector: private	-0.348	0.037	***	-2.891	0.109	***	-0.668	0.112	***
Industry2: mining and quarrying	0.352	0.121	***	0.836	0.368	**	0.441	0.131	***
Industry3: manufacturing	0.117	0.040	***	1.081	0.121	***	0.247	0.060	***
Industry4: electricity, gas and water	0.133	0.144		2.069	0.436	***	0.377	0.171	**
Industry5: construction	0.346	0.049	***	0.309	0.148	**	0.379	0.053	***
Industry6: wholesale, retail, restaurants and hotels	0.066	0.046		1.609	0.137	***	0.253	0.078	***
Industry7: transportation, storage, and communications	0.181	0.057	***	1.412	0.173	***	0.343	0.080	***
Industr8: Finance, insurance, real estate and business services	0.448	0.094	***	3.684	0.282	***	0.881	0.173	***
Industry9: Social services	0.130	0.038	***	2.097	0.114	***	0.370	0.088	***
Firm size2: 5-19 people	0.185	0.027	***	0.893	0.082	***	0.290	0.045	***

Firm size3: 20-99 people	0.253	0.032	***	1.383	0.095	***	0.415	0.063	***
Firm size4: >= 100 people	0.335	0.043	***	1.526	0.128	***	0.517	0.075	***
Province 2: Sumatera Utara (12)	0.383	0.086	***	-0.525	0.261	**	0.308	0.094	***
Province 3: Sumatera Barat (13)	0.286	0.111	***	-0.678	0.337	**	0.196	0.121	
Province 4: Riau (14)	1.036	0.157	***	-0.279	0.476		1.001	0.166	***
Province 5: Jambi (15)	0.231	0.080	***	0.504	0.244	**	0.275	0.086	***
Province 7: Lampung (18)	0.154	0.093	*	-0.153	0.282		0.133	0.098	
Province 8: Kepulauan Bangka Belitung (19)	0.564	0.077	***	0.059	0.233		0.562	0.081	***
Province 9: Kepulauan Riau (21)	0.432	0.074	***	0.015	0.225		0.428	0.078	***
Province 10: DKI Jakarta (31)	0.181	0.075	**	-0.756	0.227	***	0.090	0.084	
Province 11: Jawa Barat (32)	0.052	0.081		0.252	0.247		0.071	0.086	
Province 12: Jawa Tengah (33)	0.204	0.070	***	-0.212	0.214		0.170	0.075	**
Province 13: D I Yogyakarta (34)	0.144	0.111		-0.384	0.339		0.093	0.119	
Province 15: Banten (36)	0.788	0.333	**	1.029	1.014		0.876	0.353	**
Province 16: Bali (51)	0.339	0.076	***	-0.224	0.231		0.309	0.081	***
Province 17: Kalimantan Timur	0.919	0.440	**	-0.863	1.337		0.816	0.466	*
Province 18: Kalimantan Barat (61)	0.026	0.078		-0.618	0.236	***	-0.052	0.086	
Urban	0.009	0.026		0.946	0.078	***	0.114	0.044	***
INPRES				0.808	0.088	***			
Constant	6.569	0.135	***	9.829	0.399	***	7.794	0.425	***
Observation	6386			6386			6386		
R Squared	0.36			0.36			0.28		
Instruments:									
F-Test								84.5	***
Underidentification test (Anderson canon. corr. LM statistic):								83.991	***
Sargan statistic (overidentification test of all instruments):							(equation exactly identified)		
Endogeneity test of endogenous regressors:								10.55	***

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table VIII.5: OLS and IV Specifications without Experience and Experience Squared in 2004, All Individuals

				IV Model 1						IV Model 2					
	OLS			IV: First Stage			IV: Second Stage			IV: First Stage			IV: Second Stage		
	Coef.	SE	P>t	Coef.	SE	P>t	Coef.	SE	P>z	Coef.	SE	P>t	Coef.	SE	P>z
Years of schooling	0.087	0.004	***				-0.033	0.034					-0.034	0.034	
Sex (1=female)	-0.283	0.022	***	0.124	0.066	*	-0.262	0.024	***	0.130	0.066	**	-0.262	0.024	***
Married and cohabitate	0.169	0.027	***	-0.473	0.093	***	0.084	0.038	***	-0.541	0.083	***	0.083	0.038	**
Other (Separated, divorced and widowed)	0.246	0.056	***	-1.167	0.175	***	0.059	0.079		-1.240	0.169	***	0.057	0.080	
Religion1: Islam	0.058	0.091		-1.521	0.272	***	-0.123	0.109		-1.523	0.272	***	-0.125	0.109	
Religion2: Christian/Protestant	0.278	0.123	**	-0.047	0.369		0.265	0.130	**	-0.055	0.369		0.265	0.130	**
Religion3: Catholic	0.223	0.111	**	-0.688	0.333	**	0.139	0.120		-0.695	0.333	**	0.139	0.120	
Religion5: Buddhist	0.125	0.302		-0.771	0.906		0.026	0.321		-0.775	0.906		0.025	0.322	
Ethnicity1: Jawa	0.133	0.036	***	0.460	0.109	***	0.187	0.042	***	0.463	0.109	***	0.187	0.042	***
Ethnicity2: Sunda	0.030	0.046		0.124	0.139		0.041	0.049		0.130	0.139		0.042	0.049	
Ethnicity3: Batak	0.027	0.076		1.043	0.228	***	0.150	0.088	*	1.041	0.228	***	0.151	0.088	*
Ethnicity4: Betawi	-0.081	0.165		0.596	0.494		-0.016	0.175		0.600	0.494		-0.015	0.176	
Ethnicity5: Minang	0.158	0.072	**	0.871	0.216	***	0.264	0.082	***	0.866	0.216	***	0.265	0.082	***
Ethnicity6: Tiong Hoa	0.364	0.167	**	0.841	0.501	*	0.467	0.179	***	0.850	0.501	*	0.468	0.180	***
Status: fulltime (30 hours a week or more)	-0.467	0.029	***	-0.347	0.087	***	-0.507	0.033	***	-0.346	0.087	***	-0.507	0.033	***
Tenure	0.026	0.004	***	0.085	0.013	***	0.034	0.005	***	0.080	0.013	***	0.034	0.005	***
Tenure squared	0.000	0.000		-0.003	0.000	***	0.000	0.000	**	-0.003	0.000	***	0.000	0.000	**
Sector: private	-0.310	0.034	***	-2.297	0.099	***	-0.580	0.084	***	-2.294	0.099	***	-0.583	0.085	***
Industry2: mining and quarrying	0.372	0.087	***	0.945	0.261	***	0.483	0.097	***	0.948	0.261	***	0.484	0.098	***
Industry3: manufacturing	0.035	0.045	***	0.501	0.136	***	0.100	0.051	*	0.502	0.136	***	0.101	0.052	*
Industry4: electricity, gas and water	0.144	0.112		1.601	0.337	***	0.340	0.131	***	1.588	0.337	***	0.342	0.132	***
Industry5: construction	0.316	0.062	***	0.188	0.188		0.332	0.066	***	0.182	0.188		0.332	0.066	***
Industry6: wholesale, retail, restaurants and hotels	-0.051	0.046		1.422	0.137	***	0.126	0.070	*	1.429	0.137	***	0.128	0.070	*
Industry7: transportation, storage, and communications	0.112	0.072		1.517	0.216	***	0.292	0.092	***	1.505	0.216	***	0.293	0.092	***
Industr8: Finance, insurance, real estate and business services	0.221	0.057	***	2.793	0.167	***	0.561	0.113	***	2.794	0.167	***	0.565	0.114	***
Industry9: Social services	-0.138	0.044	***	2.828	0.128	***	0.205	0.107	*	2.828	0.128	***	0.208	0.108	*

Firm size2: 5-19 people	0.058	0.029	**	1.248	0.086	***	0.215	0.054	***	1.251	0.086	***	0.216	0.054	***
Firm size3: 20-99 people	0.233	0.031	***	1.753	0.092	***	0.452	0.070	***	1.757	0.092	***	0.454	0.071	***
Firm size4: >= 100 people	0.496	0.035	***	2.121	0.102	***	0.763	0.084	***	2.127	0.102	***	0.765	0.085	***
Province 1: Aceh	1.227	0.919		3.014	2.758		1.615	0.980	*	3.119	2.758		1.619	0.981	*
Province 2: Sumatera Utara	0.389	0.072	***	-1.105	0.215	***	0.253	0.085	***	-1.106	0.215	***	0.252	0.085	***
Province 3: Sumatera Barat	0.369	0.090	***	-0.705	0.271	***	0.276	0.099	***	-0.708	0.271	***	0.275	0.099	***
Province 4: Riau	0.724	0.131	***	-0.300	0.392		0.684	0.139	***	-0.302	0.392		0.684	0.139	***
Province 5: Jambi	0.565	0.269	**	-1.167	0.806		0.410	0.288		-1.171	0.806		0.409	0.288	
Province 6: Sumatera Selatan	0.363	0.068	***	-0.249	0.205		0.326	0.073	***	-0.256	0.205		0.326	0.073	***
Province 7: Lampung	0.151	0.077	**	-0.916	0.230	***	0.039	0.087		-0.918	0.230	***	0.038	0.087	
Province 8: Kepulauan Bangka Belitung	0.579	0.110	***	0.270	0.330		0.603	0.117	***	0.266	0.330		0.603	0.117	***
Province 9: Kepulauan Riau	0.866	0.184	***	-0.766	0.553		0.791	0.196	***	-0.775	0.553		0.791	0.197	***
Province 10: DKI Jakarta	0.679	0.058	***	-0.871	0.174	***	0.567	0.069	***	-0.874	0.174	***	0.566	0.070	***
Province 11: Jawa Barat	0.421	0.061	***	-0.840	0.184	***	0.313	0.072	***	-0.848	0.184	***	0.312	0.072	***
Province 12: Jawa Tengah	0.101	0.061	*	-1.141	0.183	***	-0.044	0.077		-1.149	0.183	***	-0.045	0.077	
Province 13: D I Yogyakarta	0.059	0.069		-0.213	0.206		0.019	0.074		-0.219	0.206		0.019	0.074	
Province 14: Jawa Timur	0.125	0.058	**	-0.933	0.175	***	0.005	0.070		-0.936	0.175	***	0.004	0.071	
Province 15: Banten	0.614	0.068	***	-1.340	0.204	***	0.454	0.085	***	-1.341	0.204	***	0.452	0.086	***
Province 16: Bali	0.596	0.095	***	-1.623	0.283	***	0.400	0.115	***	-1.623	0.283	***	0.398	0.115	***
Province 18: Kalimantan Barat	1.423	0.461	***	-3.749	1.382	***	0.937	0.507	*	-3.792	1.382	***	0.932	0.508	*
Province 19: Kalimantan Tengah	1.275	0.243	***	-0.875	0.728		1.178	0.259	***	-0.882	0.728		1.177	0.259	***
Province 20: Kalimantan Selatan	0.400	0.084	***	-0.640	0.251	**	0.321	0.092	***	-0.636	0.251	**	0.320	0.092	***
Province 21: Kalimantan Timur	0.912	0.143	***	-0.417	0.430		0.871	0.152	***	-0.438	0.430		0.871	0.152	***
Province 22: Sulawesi Selatan	0.327	0.065	***	-0.614	0.196	***	0.249	0.073	***	-0.622	0.196	***	0.249	0.073	***
Province 23: Sulawesi Tenggara	0.213	0.266		-1.094	0.800		0.068	0.285		-1.103	0.800		0.066	0.286	
Urban	0.154	0.024	***	0.764	0.073	***	0.242	0.036	***	0.757	0.073	***	0.242	0.036	***
CSAL1				0.747	0.079	***				0.783	0.076	***			
CSAL2				0.144	0.089										
Constant	8.139	0.158	***	16.199	0.444	***	10.169	0.596	***	16.285	0.441	***	10.188	0.603	***
Observation		8119		8119			8119			8119			8119		
The R-squared statistic		0.2995		0.3679			0.2074			0.3677			0.2056		

Test Results on Instruments					
F-Test	53.9	***		105.16	***
Underidentification test (Anderson canon. corr. LM statistic):	107.1	***		104.487	***
Sargan statistic (overidentification test of all instruments):	0.046			(equation exactly identified)	
Endogeneity test of endogenous regressors:	14.268	***		14.183	***
Hausman Test (Ho: difference in coefficients not systematic)	14.17	***		14.09	***

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table VIII.6: Mincer Wage Equation: IV Model by Gender, 2000

	Female						Male					
	IV: First Stage			IV: Second Stage			IV: First Stage			IV: Second Stage		
2000: Gender	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>z
Years of School				0.045	0.048					-0.081	0.062	
Married and cohabitate	-0.762	0.144	***	0.089	0.070		-0.047	0.108		0.265	0.041	***
Other (Separated, divorced and widowed)	-1.823	0.226	***	-0.060	0.133		-0.440	0.362		0.123	0.140	
Religion1: Islam	0.271	0.517		0.188	0.178		-0.776	0.411	*	-0.248	0.161	
Religion2: Christian/Protestant	1.458	0.577	**	0.509	0.208	**	0.413	0.473		0.002	0.177	
Religion3: Catholic	1.799	0.636	***	0.427	0.233	*	0.493	0.484		0.018	0.181	
Religion5: Buddhist	0.381	1.032		-0.062	0.354		-0.866	0.739		-0.343	0.281	
Ethnicity1: Jawa	-0.057	0.252		-0.080	0.087		0.103	0.173		-0.019	0.064	
Ethnicity2: Sunda	-0.048	0.304		0.005	0.104		-0.516	0.211	**	-0.212	0.085	**
Ethnicity3: Batak	0.796	0.490		-0.025	0.173		0.434	0.350		0.211	0.132	
Ethnicity4: Betawi	0.596	0.340	*	0.179	0.120		-0.060	0.250		-0.095	0.093	
Ethnicity5: Minang	1.876	0.477	***	0.229	0.187		0.524	0.349		0.085	0.133	
Ethnicity6: Tiong Hoa	0.228	0.791		0.086	0.271		0.373	0.595		0.233	0.222	
Status: fulltime (30 hours a week or more)	-0.120	0.147		-0.603	0.050	***	-0.137	0.129		-0.748	0.048	***
Tenure	0.078	0.023	***	0.040	0.008	***	0.000	0.017		0.023	0.006	***
Tenure squared	-0.002	0.001	**	-0.001	0.000	**	-0.001	0.001	*	-0.001	0.000	**
Sector: private	-3.415	0.188	***	-0.763	0.168	***	-2.560	0.132	***	-0.720	0.161	***
Industry2: mining and quarrying	1.878	1.496		1.381	0.520	***	0.615	0.382		0.404	0.146	***
Industry3: manufacturing	0.871	0.209	***	0.214	0.085	**	1.324	0.147	***	0.386	0.100	***
Industry4: electricity, gas and water	4.964	1.068	***	0.700	0.440		1.519	0.474	***	0.413	0.200	**
Industry5: construction	2.081	0.494	***	0.255	0.197		0.129	0.161		0.348	0.060	***
Industry6: wholesale, retail, restaurants and hotels	1.948	0.227	***	0.288	0.125	**	1.531	0.174	***	0.336	0.115	***
Industry7: transportation, storage, and communications	3.630	0.658	***	0.121	0.289		1.203	0.185	***	0.401	0.100	***
Industr8: Finance, insurance, real estate and business services	4.251	0.460	***	0.925	0.266	***	3.436	0.350	***	1.015	0.251	***
Industry9: Social services	2.476	0.202	***	0.349	0.141	**	1.939	0.139	***	0.525	0.129	***
Firm size2: 5-19 people	1.131	0.145	***	0.436	0.077	***	0.785	0.099	***	0.252	0.061	***
Firm size3: 20-99 people	1.586	0.162	***	0.567	0.097	***	1.305	0.118	***	0.390	0.092	***

Firm size4: >= 100 people	1.205	0.206	***	0.687	0.096	***	1.901	0.164	***	0.562	0.134	***
Province 2: Sumatera Utara (12)	-0.805	0.444	*	0.561	0.160	***	-0.360	0.317		0.217	0.121	*
Province 3: Sumatera Barat (13)	-0.878	0.575		0.140	0.204		-0.561	0.409		0.244	0.156	
Province 4: Riau (14)	0.025	0.925		1.262	0.317	***	-0.310	0.549		0.885	0.205	***
Province 5: Jambi (15)	0.533	0.439		0.295	0.151	*	0.464	0.288		0.296	0.110	***
Province 7: Lampung (18)	-0.375	0.527		0.038	0.182		-0.058	0.331		0.157	0.123	
Province 8: Kepulauan Bangka Belitung (19)	-0.055	0.397		0.754	0.136	***	0.193	0.284		0.507	0.106	***
Province 9: Kepulauan Riau (21)	0.077	0.392		0.545	0.134	***	-0.025	0.270		0.392	0.100	***
Province 10: DKI Jakarta (31)	-0.295	0.393		0.204	0.136		-0.924	0.274	***	0.001	0.117	
Province 11: Jawa Barat (32)	0.384	0.419		0.159	0.144		0.192	0.303		0.078	0.113	
Province 12: Jawa Tengah (33)	-0.210	0.377		0.338	0.130	***	-0.164	0.256		0.108	0.096	
Province 13: D I Yogyakarta (34)	-0.151	0.552		0.568	0.190	***	-0.440	0.423		-0.140	0.160	
Province 15: Banten (36)	3.159	1.844	*	1.706	0.649	***	0.241	1.196		0.650	0.444	
Province 16: Bali (51)	-0.689	0.423		0.600	0.150	***	-0.030	0.271		0.220	0.101	**
Province 17: Kalimantan Timur	-1.175	2.626		0.366	0.902		-1.015	1.536		0.895	0.574	
Province 18: Kalimantan Barat (61)	-0.430	0.403		0.180	0.141		-0.712	0.286	**	-0.201	0.115	*
Urban	0.869	0.138	***	0.044	0.061		0.932	0.094	***	0.198	0.066	***
INPRES	1.066	0.150	***				0.644	0.107	***			
Constant	8.875	0.660	***	6.550	0.521	***	10.135	0.494	***	8.788	0.692	***
Observation	2170			2170			4216			4216		
R squared (Partial R squared)	0.02			0.43			0.01			0.09		
Test Results on Instruments												
F-Test				50.75	***					36.08	***	
Underidentification test (Anderson canon. corr. LM statistic):				50.62	***					36.16	***	
Sargan statistic (overidentification test of all instruments):				(equation exactly identified)						(equation exactly identified)		
Endogeneity test of endogenous regressors:				3.52	*					10.90	***	

Source: The author’s calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table VIII.7: Mincer Wage Equation: IV Model by Gender, 2014

2014: Gender	Female						Male					
	IV: First Stage			IV: Second Stage			IV: First Stage			IV: Second Stage		
	Coef.	SE	P>t	Coef.	SE	P>z	Coef.	SE	P>t	Coef.	SE	P>z
Years of School				-0.020	0.035					-0.058	0.076	
Married and cohabitate	-1.009	0.132	***	-0.106	0.064	*	-0.221	0.106	**	0.220	0.044	***
Other (Separated, divorced and widowed)	-1.709	0.218	***	-0.104	0.110		-0.704	0.274	***	0.193	0.116	*
Religion1: Islam	-0.794	0.415	*	0.038	0.156		-1.867	0.353	***	-0.252	0.187	
Religion2: Christian/Protestant	0.598	0.554		0.462	0.205	**	-0.252	0.483		0.138	0.171	
Religion3: Catholic	-0.370	0.507		0.386	0.188	**	-0.833	0.434	*	-0.054	0.165	
Religion5: Buddhist	1.201	1.536		0.267	0.568		-1.285	1.109		-0.085	0.402	
Ethnicity1: Jawa	0.672	0.177	***	0.234	0.069	***	0.346	0.137	**	0.169	0.055	***
Ethnicity2: Sunda	0.677	0.228	***	0.098	0.087		-0.142	0.174		0.012	0.062	
Ethnicity3: Batak	1.547	0.359	***	0.182	0.142		0.657	0.291	**	0.111	0.114	
Ethnicity4: Betawi	1.672	0.747	**	-0.070	0.281		0.018	0.661		0.015	0.232	
Ethnicity5: Minang	1.375	0.343	***	0.321	0.135	**	0.579	0.276	**	0.166	0.107	
Ethnicity6: Tiong Hoa	2.046	0.841	**	0.572	0.320	*	0.235	0.620		0.368	0.218	*
Status: fulltime (30 hours a week or more)	-0.260	0.121	**	-0.390	0.045	***	-0.443	0.124	***	-0.687	0.055	***
Tenure	0.145	0.020	***	0.048	0.009	***	0.020	0.016		0.022	0.006	***
Tenure squared	-0.004	0.001	***	-0.001	0.000	*	-0.001	0.001	**	0.000	0.000	
Sector: private	-2.354	0.151	***	-0.609	0.098	***	-2.233	0.131	***	-0.588	0.175	***
Industry2: mining and quarrying	3.378	1.035	***	0.856	0.402	**	0.704	0.276	***	0.528	0.110	***
Industry3: manufacturing	0.261	0.240		-0.166	0.090	*	0.747	0.164	***	0.260	0.082	***
Industry4: electricity, gas and water	3.007	1.370	**	0.074	0.516		1.382	0.351	***	0.454	0.163	***
Industry5: construction	2.850	0.636	***	0.765	0.256	***	-0.196	0.204		0.298	0.074	***
Industry6: wholesale, retail, restaurants and hotels	1.664	0.238	***	-0.064	0.110		1.395	0.168	***	0.269	0.122	**
Industry7: transportation, storage, and communications	1.153	0.651	*	0.289	0.245		1.368	0.233	***	0.359	0.131	***

Industr8: Finance, insurance, real estate and business services	3.243	0.313	***	0.501	0.168	***	2.537	0.196	***	0.635	0.206	***
Industry9: Social services	3.458	0.225	***	0.098	0.153		2.388	0.157	***	0.286	0.190	
Firm size2: 5-19 people	1.483	0.133	***	0.183	0.075	**	1.012	0.111	***	0.220	0.088	**
Firm size3: 20-99 people	2.030	0.146	***	0.506	0.093	***	1.567	0.116	***	0.420	0.129	***
Firm size4: >= 100 people	2.007	0.166	***				2.165	0.128	***	0.741	0.174	***
Province 1: Aceh				0.848	0.100	***	2.869	2.744		1.631	0.988	***
Province 2: Sumatera Utara	-1.195	0.345	***	0.274	0.134	**	-1.073	0.270	***	0.246	0.126	*
Province 3: Sumatera Barat	-0.836	0.427	**	0.484	0.161	***	-0.754	0.346	**	0.181	0.136	
Province 4: Riau	-1.569	0.794	**	0.721	0.298	**	-0.075	0.451		0.708	0.158	***
Province 5: Jambi	2.207	1.565		0.942	0.582		-2.402	0.928	***	0.084	0.376	
Province 6: Sumatera Selatan	-0.128	0.349		0.484	0.129	***	-0.462	0.250	*	0.229	0.096	**
Province 7: Lampung	-1.012	0.389	***	-0.020	0.148		-0.973	0.281	***	0.028	0.124	
Province 8: Kepulauan Bangka Belitung	1.090	0.522	**	0.802	0.195	***	-0.205	0.418		0.480	0.148	***
Province 9: Kepulauan Riau	-0.837	1.133		1.036	0.419	**	-0.748	0.629		0.744	0.226	***
Province 10: DKI Jakarta	-0.694	0.286	**	0.695	0.109	***	-1.031	0.216	***	0.478	0.111	***
Province 11: Jawa Barat	-0.906	0.296	***	0.412	0.114	***	-0.826	0.230	***	0.252	0.104	**
Province 12: Jawa Tengah	-0.688	0.289	**	0.015	0.110		-1.440	0.233	***	-0.089	0.139	
Province 13: D I Yogyakarta	0.061	0.325		0.157	0.120		-0.282	0.263		-0.043	0.096	
Province 14: Jawa Timur	-0.765	0.279	***	0.063	0.107		-0.981	0.220	***	-0.018	0.110	
Province 15: Banten	-1.231	0.319	***	0.738	0.125	***	-1.355	0.261	***	0.249	0.138	*
Province 16: Bali	-1.043	0.437	**	0.729	0.166	***	-1.850	0.365	***	0.189	0.190	
Province 18: Kalimantan Barat	-2.962	1.912		1.308	0.715	*	-4.448	1.937	**	0.538	0.765	
Province 19: Kalimantan Tengah	-1.328	1.232		0.737	0.456		-0.900	0.887		1.354	0.319	***
Province 20: Kalimantan Selatan	-0.517	0.397		0.265	0.147	**	-0.752	0.320	**	0.373	0.127	***
Province 21: Kalimantan Timur	-0.870	0.752		0.805	0.278	***	-0.299	0.517		0.941	0.182	***
Province 22: Sulawesi Selatan	-0.519	0.303	*	0.137	0.113		-0.683	0.251	***	0.345	0.103	***
Province 23: Sulawesi Tenggara	1.199	1.553		-0.195	0.574		-1.638	0.921	*	0.098	0.347	
Urban	0.728	0.113	***	0.272	0.048	***	0.783	0.093	***	0.242	0.067	***

CSAL1	1.294	0.123	***				0.441	0.096	***			
Constant	14.765	0.672	***	9.661	0.604	***	17.203	0.575	***	10.706	1.355	***
Observation	3173			3173			4946			4946		
The R-squared statistic (Partial R-squared)	0.0343			0.263			0.0043			0.146		
Test Results on Instruments												
F-Test				110.81 ***						21.12 ***		
Underidentification test (Anderson canon. corr. LM statistic):				108.763 ***						21.257 ***		
Sargan statistic (overidentification test of all instruments):				(equation exactly identified)						(equation exactly identified)		
Endogeneity test of endogenous regressors:				11.846 ***						3.995 **		

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table VIII.8: Mincer Wage Equation: IV Model by Sector, 2000

	Private						Public					
	IV: First Stage			IV: Second Stage			IV: First Stage			IV: Second Stage		
2000: Sector	Coef.	SE	P>t	Coef.	SE	P>z	Coef.	SE	P>t	Coef.	SE	P>z
Years of School				0.024	0.035					-0.176	0.224	
Female	-0.533	0.082	***	-0.453	0.034	***	0.885	0.168	***	0.185	0.213	
Married and cohabitate	-0.266	0.091	***	0.187	0.035	***	-0.239	0.294		0.236	0.150	
Other (Separated, divorced and widowed)	-1.202	0.197	***	-0.035	0.087		-1.523	0.537	***	-0.178	0.446	
Religion1: Islam	-0.367	0.382		-0.072	0.132		-0.628	0.578		-0.130	0.289	
Religion2: Christian/Protestant	1.101	0.435	**	0.213	0.153		-0.050	0.658		-0.037	0.280	
Religion3: Catholic	1.789	0.454	***	0.287	0.167	*	-1.604	0.709	**	-0.703	0.472	
Religion5: Buddhist	-0.464	0.639		-0.209	0.220		3.118	2.606		0.984	1.303	
Ethnicity1: Jawa	0.114	0.160		-0.050	0.055		0.052	0.306		0.055	0.130	
Ethnicity2: Sunda	-0.321	0.194	*	-0.087	0.067		-0.181	0.393		-0.152	0.172	
Ethnicity3: Batak	0.836	0.318	***	0.147	0.113		1.539	0.695	**	0.586	0.451	
Ethnicity4: Betawi	0.305	0.218		0.041	0.075		0.223	0.579		-0.245	0.252	
Ethnicity5: Minang	1.318	0.318	***	0.115	0.119		0.277	0.593		0.238	0.256	
Ethnicity6: Tiong Hoa	0.182	0.483		0.092	0.165							
Status: fulltime (30 hours a week or more)	-0.124	0.107		-0.703	0.037	***	-0.103	0.219		-0.495	0.096	***
Tenure	0.016	0.015		0.029	0.005	***	0.123	0.033	***	0.072	0.026	***
Tenure squared	-0.001	0.001	**	-0.001	0.000	***	-0.004	0.001	***	-0.001	0.001	*
Industry2: mining and quarrying	0.676	0.374	*	0.358	0.130	***	2.861	1.785		1.744	0.976	*
Industry3: manufacturing	1.032	0.126	***	0.194	0.057	***	0.324	0.482		0.171	0.219	
Industry4: electricity, gas and water	2.870	0.547	***	0.441	0.213	**	0.531	0.721		0.128	0.335	
Industry5: construction	0.250	0.153		0.367	0.053	***	0.259	0.606		0.201	0.263	
Industry6: wholesale, retail, restaurants and hotels	1.513	0.142	***	0.189	0.073	***	1.079	0.664		0.151	0.371	
Industry7: transportation, storage, and communications	1.324	0.183	***	0.287	0.078	***	1.138	0.501	**	0.523	0.334	
Industr8: Finance, insurance, real estate and business services	3.843	0.309	***	0.815	0.173	***	2.693	0.650	***	1.105	0.683	
Industry9: Social services	1.932	0.123	***	0.257	0.080	***	1.925	0.318	***	0.726	0.451	
Firm size2: 5-19 people	0.826	0.088	***	0.244	0.043	***	0.913	0.230	***	0.451	0.221	**
Firm size3: 20-99 people	1.411	0.104	***	0.384	0.062	***	1.355	0.247	***	0.599	0.315	*

Firm size4: >= 100 people	1.729	0.138	***	0.548	0.079	***	0.794	0.347	**	0.339	0.229	
Province 2: Sumatera Utara (12)	-0.192	0.288		0.462	0.099	***	-2.975	0.653	***	-0.868	0.723	
Province 3: Sumatera Barat (13)	-0.616	0.379		0.292	0.133	**	-0.892	0.681		-0.216	0.351	
Province 4: Riau (14)	-0.273	0.503		1.161	0.173	***	-1.255	1.520		-0.430	0.701	
Province 5: Jambi (15)	0.759	0.272	***	0.347	0.095	***	-0.409	0.513		-0.097	0.238	
Province 7: Lampung (18)	0.027	0.309		0.208	0.106	**	-0.756	0.689		-0.125	0.335	
Province 8: Kepulauan Bangka Belitung (19)	0.131	0.261		0.661	0.089	***	-0.441	0.508		0.095	0.235	
Province 9: Kepulauan Riau (21)	0.062	0.255		0.540	0.087	***	0.012	0.456		0.054	0.193	
Province 10: DKI Jakarta (31)	-0.721	0.256	***	0.208	0.092	**	-0.641	0.473		-0.323	0.242	
Province 11: Jawa Barat (32)	0.260	0.281		0.095	0.096		0.251	0.490		0.070	0.214	
Province 12: Jawa Tengah (33)	-0.160	0.243		0.256	0.084	***	0.029	0.433		0.004	0.184	
Province 13: D I Yogyakarta (34)	-0.032	0.396		0.178	0.136		-1.154	0.609	*	-0.286	0.373	
Province 15: Banten (36)	3.238	1.190	***	1.267	0.420	***	-3.496	1.807	*	-0.664	1.096	
Province 16: Bali (51)	0.056	0.270		0.408	0.093	***	-0.936	0.412	**	-0.134	0.273	
Province 17: Kalimantan Timur	-0.941	1.332		0.882	0.458	*						
Province 18: Kalimantan Barat (61)	-0.384	0.282		0.087	0.098		-0.953	0.404	**	-0.470	0.278	*
Urban	1.151	0.086	***	0.112	0.049	**	0.116	0.172		0.046	0.076	
Inpres	0.917	0.094	***				0.431	0.228	*	9.332	2.434	***
Constant	6.723	0.448	***	6.942	0.311	***	10.251	0.791	***			
Observation	5326						1060			1060		
R squared (Partial R squared)	0.0176						0.0035			-0.2718		
Test Results on Instruments												
F-Test				94.7	***					3.59	*	
Underidentification test (Anderson canon. corr. LM statistic):				93.821	***					3.724	*	
Sargan statistic (overidentification test of all instruments):				(equation exactly identified)						(equation exactly identified)		
Endogeneity test of endogenous regressors:				5.242	**					3.469	*	

Source: The author’s calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** Significance level at 1 per cent.

Table VIII.9: Mincer Wage Equation: IV Model by Sector, 2014

	Private						Public					
	IV: First Stage			IV: Second Stage			IV: First Stage			IV: Second Stage		
2014	**Coef.**	**SE**	**P>t**	**Coef.**	**SE**	**P>z**	**Coef.**	**SE**	**P>t**	**Coef.**	**SE**	**P>z**
Years of School				0.067	0.031	**				-1.599	0.980	
Female	-0.028	0.075		-0.299	0.024	***	0.825	0.132	***	1.262	0.868	
Married and cohabitate	-0.533	0.090	***	0.102	0.036	***	-0.075	0.225		0.383	0.407	
Other (Separated, divorced and widowed)	-1.270	0.185	***	0.133	0.076	*	-0.546	0.407		-0.171	0.944	
Religion1: Islam	-1.541	0.309	***	-0.009	0.108		-1.377	0.532	***	-2.021	1.661	
Religion2: Christian/Protestant	0.318	0.422		0.329	0.134	**	-1.142	0.697		-1.670	1.706	
Religion3: Catholic	-0.455	0.379		0.256	0.120	**	-1.356	0.649	**	-2.179	1.774	
Religion5: Buddhist	-0.753	0.960		0.028	0.304		0.967	2.374		3.525	4.243	
Ethnicity1: Jawa	0.523	0.122	***	0.107	0.041	***	0.079	0.240		0.376	0.424	
Ethnicity2: Sunda	0.098	0.152		-0.002	0.048		-0.037	0.344		0.027	0.602	
Ethnicity3: Batak	1.089	0.257	***	-0.012	0.088		0.965	0.463	**	1.848	1.233	
Ethnicity4: Betawi	1.054	0.586	*	-0.191	0.187		-0.104	0.836		0.049	1.460	
Ethnicity5: Minang	1.071	0.251	***	0.142	0.086	*	0.183	0.393		0.592	0.706	
Ethnicity6: Tiong Hoa	0.684	0.519		0.302	0.165	*						
Status: fulltime (30 hours a week or more)	-0.312	0.102	***	-0.523	0.034	***	-0.405	0.151	***	-1.032	0.476	**
Tenure	0.097	0.015	***	0.041	0.005	***	0.068	0.026	***	0.123	0.074	*
Tenure squared	-0.004	0.001	***	-0.001	0.000	***	-0.002	0.001	**	-0.003	0.002	
Industry2: mining and quarrying	0.954	0.285	***	0.380	0.094	***	0.274	0.663		0.881	1.177	
Industry3: manufacturing	0.468	0.146	***	0.064	0.049		-0.506	0.657		-1.288	1.255	
Industry4: electricity, gas and water	1.451	0.371	***	0.255	0.125	**	1.434	0.767	*	1.998	1.996	
Industry5: construction	0.216	0.196		0.334	0.062	***	-2.542	1.190	**	-4.286	3.203	
Industry6: wholesale, retail, restaurants and hotels	1.440	0.146	***	-0.008	0.065		-1.271	0.948		-2.562	2.035	
Industry7: transportation, storage, and communications	1.429	0.227	***	0.129	0.084		1.638	0.883	*	3.503	2.248	
Industr8: Finance, insurance, real estate and business services	2.707	0.180	***	0.288	0.101	***	2.610	0.478	***	4.606	2.714	*
Industry9: Social services	2.857	0.140	***	-0.090	0.099		2.098	0.331	***	3.440	2.130	
Firm size2: 5-19 people	1.169	0.094	***	0.072	0.048		0.998	0.244	***	1.837	1.068	*

Firm size3: 20-99 people	1.772	0.101	***	0.253	0.065	***	1.169	0.243	***	2.306	1.233	*
Firm size4: >= 100 people	2.333	0.111	***	0.552	0.083	***	0.906	0.272	***	2.088	1.031	**
Province 1: Aceh	2.838	2.819		1.183	0.894							
Province 2: Sumatera Utara	-0.974	0.251	***	0.349	0.085	***	-1.350	0.398	***	-1.762	1.475	
Province 3: Sumatera Barat	-0.913	0.324	***	0.282	0.107	***	-0.265	0.456		-0.042	0.829	
Province 4: Riau	-0.381	0.442		0.590	0.140	***	-0.177	0.807		0.928	1.412	
Province 5: Jambi	-1.218	0.828		0.473	0.264	*				-0.084	0.668	
Province 6: Sumatera Selatan	-0.214	0.243		0.313	0.077	***	-0.249	0.360		-0.399	0.881	
Province 7: Lampung	-1.005	0.264	***	0.075	0.089		-0.343	0.464		1.500	1.030	
Province 8: Kepulauan Bangka Belitung	0.324	0.407		0.596	0.128	***	0.517	0.495		-1.846	2.452	
Province 9: Kepulauan Riau	-0.405	0.627		0.847	0.198	***	-1.754	1.076		-0.385	0.951	
Province 10: DKI Jakarta	-0.926	0.201	***	0.619	0.071	***	-0.651	0.415		-0.089	0.718	
Province 11: Jawa Barat	-0.913	0.214	***	0.362	0.074	***	-0.309	0.372		0.265	0.620	
Province 12: Jawa Tengah	-1.308	0.214	***	0.011	0.080		0.001	0.356		0.653	0.720	
Province 13: D I Yogyakarta	-0.299	0.240		-0.033	0.077		0.258	0.396		-1.101	0.881	
Province 14: Jawa Timur	-0.945	0.206	***	0.087	0.072		-0.712	0.320	**	-2.472	1.847	
Province 15: Banten	-1.441	0.230	***	0.553	0.085	***	-1.488	0.627	**	-2.031	2.027	
Province 16: Bali	-1.603	0.324	***	0.454	0.113	***	-1.775	0.558	***			
Province 18: Kalimantan Barat	-3.781	1.415	***	1.299	0.463	***						
Province 19: Kalimantan Tengah	-0.750	0.870		1.396	0.275	***	-0.477	1.199		-0.113	2.131	
Province 20: Kalimantan Selatan	-0.933	0.299	***	0.297	0.099	***	-0.043	0.428		0.458	0.747	
Province 21: Kalimantan Timur	-0.684	0.478		0.922	0.152	***	-0.667	1.058		-0.526	1.921	
Province 22: Sulawesi Selatan	-0.658	0.244	***	0.358	0.080	***	-0.438	0.285		-0.497	0.664	
Province 23: Sulawesi Tenggara	0.169	1.067		0.026	0.336		-2.044	1.047	*	-2.824	2.681	
Urban	0.887	0.083	***	0.132	0.037	***	0.381	0.140	***	0.936	0.431	**
CSAL1	0.882	0.086	***				0.282	0.159	*			
Constant	9.269	0.380	***	7.628	0.331	***	12.588	0.712	***	28.089	12.695	**
Observation	6728			6728			1391			1391		
The R-squared stat. (Partial R-squared)	0.0157			0.2622			0.0023			-10.2838		
Test Results on Instruments												
F-Test				106.2		***				3.16		*
Underidentification test (Anderson canon. corr. LM statistic):				105.35		***				3.267		*
Sargan statistic (overidentification test of all instruments):				(equation exactly identified)						(equation exactly identified)		
Endogeneity test of endogenous regressors:				0.181						51.609		***

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Appendix IX: IV with Smoking as the Instrument

Table IX.1: IV with Smoking as Instrument, 2000

	2000: All Individuals								
	OLS: Wages			IV: First Stage			IV: Wage equation		
	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>t
Years of Schooling	0.120	0.004	***				0.102	0.059	*
Experience	0.033	0.005	***	-0.294	0.014	***	0.028	0.018	
Experience squared	-0.001	0.000	***	0.003	0.000	***	-0.001	0.000	**
Sex (1=female)	-0.328	0.024	***	-0.776	0.090	***	-0.336	0.036	***
Married and cohabitate	0.099	0.033	***	1.416	0.095	***	0.124	0.089	
Other (Separated, divorced and widowed)	-0.052	0.064		0.898	0.185	***	-0.036	0.083	
Religion1: Islam	-0.009	0.106		-0.340	0.306		-0.015	0.108	
Religion2: Christian/Protestant	0.083	0.121		0.993	0.348	***	0.101	0.134	
Religion3: Catholic	0.052	0.127		1.127	0.367	***	0.072	0.144	
Religion5: Buddhist	-0.186	0.199		0.110	0.573		-0.181	0.199	
Ethnicity1: Jawa	-0.043	0.047		0.048	0.135		-0.042	0.047	
Ethnicity2: Sunda	-0.061	0.057		-0.165	0.165		-0.064	0.058	
Ethnicity3: Batak	0.064	0.094		0.626	0.271	**	0.074	0.100	
Ethnicity4: Betawi	0.010	0.066		0.221	0.191		0.013	0.067	
Ethnicity5: Minang	0.052	0.093		1.012	0.268	***	0.070	0.110	
Ethnicity6: Tiong Hoa	0.124	0.157		0.394	0.453		0.130	0.158	
Status: fulltime (30 hours a week or more)	-0.640	0.032	***	-0.186	0.092	**	-0.643	0.034	***
Tenure	0.022	0.005	***	0.115	0.013	***	0.024	0.008	***
Tenure squared	0.000	0.000		-0.002	0.000	***	0.000	0.000	
Sector: private	-0.303	0.038	***	-2.839	0.102	***	-0.354	0.171	**
Industry2: mining and quarrying	0.363	0.120	***	0.635	0.347	*	0.374	0.125	***
Industry3: manufacturing	0.126	0.040	***	0.736	0.114	***	0.139	0.060	**
Industry4: electricity, gas and water	0.139	0.144		1.767	0.415	***	0.171	0.178	
Industry5: construction	0.337	0.049	***	0.363	0.140	***	0.343	0.053	***
Industry6: wholesale, retail, restaurants and hotels	0.079	0.045	*	1.170	0.130	***	0.100	0.084	
Industry7: transportation, storage, and communications	0.173	0.057	***	1.234	0.163	***	0.194	0.091	**
Industry8: Finance, insurance, real estate and business services	0.418	0.093	***	3.288	0.266	***	0.476	0.214	**
Industry9: Social services	0.113	0.038	***	1.865	0.108	***	0.146	0.118	
Firm size2: 5-19 people	0.183	0.027	***	0.734	0.077	***	0.195	0.051	***
Firm size3: 20-99 people	0.242	0.032	***	1.200	0.090	***	0.264	0.078	***
Firm size4: >= 100 people	0.338	0.042	***	1.139	0.122	***	0.359	0.080	***

Province2: Sumatera Utara (12)	0.368	0.085	***	-0.128	0.246		0.365	0.086	***
Province3: Sumatera Barat (13)	0.253	0.110	**	-0.122	0.318		0.250	0.111	**
Province4: Riau (14)	1.054	0.156	***	-0.331	0.448		1.047	0.157	***
Province5: Sumatera Selatan (16)	0.202	0.080	**	0.817	0.229	***	0.216	0.092	**
Province7: Lampung (18)	0.153	0.092	*	-0.024	0.265		0.152	0.092	*
Province8: DKI Jakarta (31)	0.542	0.076	***	0.399	0.219	*	0.549	0.079	***
Province9: Jawa Barat (32)	0.419	0.074	***	0.178	0.212		0.422	0.074	***
Province10: Jawa Tengah (33)	0.167	0.074	**	-0.386	0.214	*	0.160	0.078	**
Province11: D I Yogyakarta (34)	0.023	0.081		0.631	0.233	***	0.033	0.088	
Province12: Jawa Timur (35)	0.180	0.070	***	0.146	0.202		0.183	0.070	***
Province13: Bali (51)	0.150	0.111		-0.325	0.319		0.145	0.112	
Province15: Kalimantan Tengah (62)	0.753	0.332	**	1.385	0.954		0.776	0.339	**
Province16: Kalimantan Selatan (63)	0.336	0.075	***	-0.044	0.217		0.335	0.075	***
Province17: Kalimantan Timur (64)	0.911	0.437	**	-0.720	1.258		0.900	0.437	**
Province18: Sulawesi Selatan (73)	0.010	0.077		-0.223	0.222		0.006	0.078	
Urban	-0.007	0.026		1.014	0.073	***	0.011	0.065	
Smoke				-0.501	0.085	***			
Constants	6.294	0.139	***	12.546	0.375	***	6.509	0.727	***
Observation	6383			6383			6383		
R Squared	0.371			0.006			0.370		

Test Results on Instruments		
Quality (F Test)	35.02	***
Validity (Sargan Test)	exactly identified	
Underidentification test (Anderson canon. corr. LM statistic):	35.09	0.00
Endogeneity test (years of schooling)	0.09	0.7627

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table IX.2: IV with Smoking as Instrument, 2014

Log of real hourly wage	2014: All Individuals								
	OLS: Wages			IV: First Stage			IV: Wage equation		
	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>t
Years of schooling	0.099	0.004	***				0.088	0.044	**
Experience	0.020	0.004	***	-0.147	0.011	***	0.018	0.008	**
Experience squared	0.000	0.000	***	0.000	0.000		0.000	0.000	***
Sex (1=female)	-0.268	0.022	***	-0.524	0.081	***	-0.269	0.022	***
Married and cohabitate	0.055	0.033	*	0.807	0.087	***	0.063	0.048	
Other (Separated, divorced and widowed)	0.109	0.059	*	0.843	0.160	***	0.117	0.069	*
Religion1: Islam	0.056	0.091		-1.234	0.243	***	0.043	0.105	
Religion2: Christian/Protestant	0.250	0.123	**	0.332	0.330		0.254	0.123	**
Religion3: Catholic	0.211	0.111	*	-0.472	0.298		0.207	0.112	*
Religion5: Buddhist	0.102	0.301		-0.090	0.810		0.100	0.300	
Ethnicity1: Jawa	0.129	0.036	***	0.407	0.098	***	0.134	0.041	***
Ethnicity2: Sunda	0.024	0.046		0.168	0.125		0.025	0.047	
Ethnicity3: Batak	0.021	0.076		0.956	0.204	***	0.031	0.086	
Ethnicity4: Betawi	-0.086	0.164		0.694	0.441		-0.079	0.167	
Ethnicity5: Minang	0.152	0.072	**	0.770	0.193	***	0.160	0.079	**
Ethnicity6: Tiong Hoa	0.366	0.167	**	0.544	0.448		0.373	0.169	**
Status: fulltime (30 hours a week or more)	-0.461	0.029	***	-0.375	0.078	***	-0.465	0.033	***
Tenure	0.019	0.004	***	0.115	0.012	***	0.021	0.007	***
Tenure squared	0.000	0.000		-0.001	0.000	***	0.000	0.000	
Sector: private	-0.288	0.034	***	-2.106	0.089	***	-0.310	0.099	***
Industry2: mining and quarrying	0.372	0.087	***	0.776	0.233	***	0.381	0.093	***
Industry3: manufacturing	0.041	0.045		0.217	0.122	*	0.044	0.047	
Industry4: electricity, gas and water	0.139	0.112		1.291	0.301	***	0.152	0.125	
Industry5: construction	0.306	0.062	***	0.295	0.168	*	0.309	0.064	***
Industry6: wholesale, retail, restaurants and hotels	-0.037	0.046		0.886	0.123	***	-0.027	0.061	
Industry7: transportation, storage, and communications	0.102	0.072		1.304	0.193	***	0.115	0.092	
Industr8: Finance, insurance, real estate and business services	0.229	0.057	***	2.144	0.150	***	0.252	0.111	**
Industry9: Social services	-0.137	0.044	***	2.238	0.115	***	-0.113	0.109	
Firm size2: 5-19 people	0.064	0.029	**	0.894	0.078	***	0.074	0.049	
Firm size3: 20-99 people	0.240	0.031	***	1.269	0.083	***	0.253	0.065	***
Firm size4: >= 100 people	0.514	0.035	***	1.438	0.092	***	0.529	0.073	***
Province 1: Aceh	1.318	0.917		2.043	2.465		1.337	0.918	
Province 2: Sumatera Utara	0.384	0.072	***	-0.773	0.192	***	0.376	0.080	***
Province 3: Sumatera Barat	0.355	0.090	***	-0.378	0.242		0.351	0.092	***

Province 4: Riau	0.722	0.130	***	-0.318	0.351		0.719	0.131	***
Province 5: Jambi	0.545	0.268	**	-0.783	0.720		0.537	0.269	**
Province 6: Sumatera Selatan	0.349	0.068	***	0.014	0.183		0.350	0.068	***
Province 7: Lampung	0.149	0.077	*	-0.698	0.206	***	0.142	0.082	*
Province 8: Kepulauan Bangka Belitung	0.563	0.110	***	0.530	0.297	*	0.568	0.113	***
Province 9: Kepulauan Riau	0.888	0.184	***	-0.785	0.494		0.880	0.186	***
Province 10: DKI Jakarta	0.670	0.058	***	-0.511	0.156	***	0.665	0.062	***
Province 11: Jawa Barat	0.409	0.061	***	-0.539	0.165	***	0.404	0.065	***
Province 12: Jawa Tengah	0.091	0.061		-0.805	0.164	***	0.083	0.070	
Province 13: D I Yogyakarta	0.038	0.069		0.100	0.185		0.039	0.069	
Province 14: Jawa Timur	0.116	0.058	**	-0.588	0.156	***	0.110	0.063	*
Province 15: Banten	0.611	0.068	***	-1.102	0.183	***	0.600	0.083	***
Province 16: Bali	0.592	0.094	***	-1.303	0.253	***	0.579	0.109	***
Province 18: Kalimantan Barat	1.350	0.460	***	-2.168	1.236	*	1.327	0.468	***
Province 19: Kalimantan Tengah	1.267	0.242	***	-0.734	0.651		1.260	0.243	***
Province 20: Kalimantan Selatan	0.402	0.084	***	-0.570	0.225	**	0.396	0.087	***
Province 21: Kalimantan Timur	0.911	0.143	***	-0.394	0.384		0.907	0.143	***
Province 22: Sulawesi Selatan	0.319	0.065	***	-0.432	0.175	**	0.315	0.068	***
Province 23: Sulawesi Tenggara	-0.011	0.277		-0.968	0.746		-0.020	0.279	
Urban	0.137	0.024	***	0.783	0.065	***	0.145	0.042	***
Smoke				-0.662	0.078	***			
Constant	7.857	0.163	***	18.149	0.396	***	8.043	0.790	***
Observation	8086			8086			8086		
The R-squared statistic	0.30			0.01			0.30		
Test Results on Instruments Quality									
F test							71.41	***	
Validity (Sargan Test)							71.26	***	
Underidentification test (Anderson canon. corr. LM statistic):							exactly identified		
Endogeneity test (years of schooling)							0.06	0.8104	

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Appendix X: Mincer Wage Equation: Two-Step Heckman by Gender

Table X.1: Mincer Wage Equation: Two-Step Heckman by Gender, 2000

	Female						Male					
	Probit			OLS			Probit			OLS		
2000: Gender	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>t
HH Size	-0.043	0.014	***				-0.013	0.010				
Employment	(omitted)						(omitted)					
Years of Schooling	0.142	0.015	***	0.140	0.008	***	0.097	0.011	***	0.106	0.005	***
Experience	0.062	0.017	***	0.020	0.008	**	0.050	0.013	***	0.042	0.006	***
Experience squared	-0.001	0.000	***	-0.000	0.000		-0.001	0.000	***	-0.001	0.000	***
Sex (1=female)	(omitted)			(omitted)			(omitted)			(omitted)		
Married and cohabitate	-0.092	0.128		0.091	0.055	*	-0.023	0.096		0.121	0.042	***
Other (Separated, divorced and widowed)	0.047	0.174		0.020	0.083		-0.092	0.225		0.048	0.120	
Religion1: Islam	-0.161	0.344		0.157	0.173		0.050	0.222		-0.096	0.133	
Religion2: Christian/Protestant	-0.139	0.389		0.359	0.193	*	0.240	0.274		-0.085	0.153	
Religion3: Catholic	-0.825	0.426	*	0.249	0.213		-0.049	0.292		-0.071	0.157	
Religion5: Buddhist	0.555	0.636		-0.103	0.346		0.776	0.447	*	-0.247	0.240	
Ethnicity1: Jawa	-0.124	0.164		-0.069	0.085		0.092	0.115		-0.044	0.056	
Ethnicity2: Sunda	0.264	0.217		0.013	0.102		0.207	0.153		-0.134	0.069	*
Ethnicity3: Batak	-0.177	0.314		-0.108	0.164		-0.293	0.236		0.134	0.113	
Ethnicity4: Betawi	0.223	0.275		0.142	0.114		-0.093	0.193		-0.093	0.081	
Ethnicity5: Minang	0.083	0.345		0.052	0.160		-0.170	0.239		0.005	0.113	
Ethnicity6: Tiong Hoa	-1.060	0.455	**	0.063	0.266		-1.797	0.353	***	0.227	0.194	
Status: fulltime (30 hours a week or more)	0.572	0.087	***	-0.575	0.050	***	0.340	0.073	***	-0.744	0.042	***
Tenure	0.079	0.015	***	0.028	0.008	***	0.070	0.011	***	0.011	0.006	**
Tenure squared	-0.002	0.001	***	-0.000	0.000		-0.002	0.000	***	-0.000	0.000	
Sector: private	3.734	0.109	***	-0.410	0.104	***	3.728	0.080	***	-0.363	0.070	***
Industry2: mining and quarrying	-0.547	0.928		1.204	0.499	**	0.407	0.323		0.301	0.124	**
Industry3: manufacturing	0.201	0.139		0.155	0.071	**	0.275	0.108	**	0.135	0.048	***
Industry4: electricity, gas and water	(omitted)			(omitted)			1.185	0.328	***	0.110	0.154	
Industry5: construction	0.671	0.429		0.063	0.166		0.193	0.118		0.301	0.053	***
Industry6: wholesale, retail, restaurants and hotels	-0.087	0.131		0.139	0.077	*	-0.052	0.102		0.064	0.057	

Industry7: transportation, storage, and communications	0.887	0.439	**	-0.163	0.222		0.379	0.111	***	0.156	0.060	***
Industry8: Finance, insurance, real estate and business services	0.661	0.489		0.528	0.157	***	2.229	0.517	***	0.328	0.115	***
Industry9: Social services	1.195	0.123	***	0.138	0.075	*	1.327	0.085	***	0.121	0.048	**
Firm size2: 5-19 people	1.045	0.094	***	0.345	0.053	***	0.960	0.068	***	0.079	0.034	**
Firm size3: 20-99 people	1.387	0.142	***	0.431	0.060	***	1.485	0.100	***	0.108	0.041	***
Firm size4: >= 100 people	1.554	0.255	***	0.590	0.072	***	1.543	0.172	***	0.184	0.055	***
Province1: Aceh	(omitted)			(omitted)			(omitted)			(omitted)		
Province2: Sumatera Utara	0.776	5.356		0.119	0.866		-3.722	260.959		-0.727	0.502	
Province3: Sumatera Barat	0.581	5.366		-0.295	0.882		-3.604	260.959		-0.705	0.509	
Province4: Riau	-0.027	5.436		0.796	0.908		-2.811	260.959		-0.087	0.522	
Province5: Sumatera Selatan	0.355	5.359		-0.281	0.877		-3.985	260.959		-0.799	0.499	
Province6: Bengkulu	(omitted)			(omitted)			(omitted)			(omitted)		
Province7: Lampung	0.507	5.360		-0.399	0.880		-3.864	260.959		-0.840	0.503	*
Province8: DKI Jakarta	0.377	5.357		0.230	0.870		-3.915	260.959		-0.541	0.499	
Province9: Jawa Barat	0.453	5.357		0.012	0.871		-3.752	260.959		-0.616	0.499	
Province10: Jawa Tengah	0.727	5.357		-0.291	0.871		-3.818	260.959		-0.860	0.499	*
Province11: D I Yogyakarta	0.782	5.358		-0.401	0.872		-3.898	260.959		-0.990	0.501	**
Province12: Jawa Timur	0.763	5.357		-0.174	0.871		-3.869	260.959		-0.887	0.498	*
Province13: Bali	0.702	5.365		0.070	0.884		-3.414	260.959		-1.066	0.510	**
Province15: Kalimantan Tengah	0.820	5.358		-0.512	0.876		-3.459	260.959		-1.022	0.498	**
Province16: Kalimantan Selatan	(omitted)			(omitted)			-1.730	260.962		-0.462	0.625	
Province17: Kalimantan Timur	1.295	5.358		0.142	0.877		-3.258	260.959		-0.806	0.498	
Province18: Sulawesi Selatan	0.381	5.357		-0.302	0.877		-3.437	260.959		-1.112	0.499	**
Province19: Sulawesi Tenggara	(omitted)			(omitted)			(omitted)			(omitted)		
Urban	0.017	0.092		-0.050	0.047		0.022	0.065		0.020	0.031	
IMR				0.046	0.102					-0.180	0.070	***
Constants	-5.099	5.369		5.952	0.912	***	-0.511	260.959		7.769	0.527	***
Number of jobs (Uncensored jobs)	5438			2162			7861			4216		
The R-squared statistic				0.47						0.32		
LR chi2(47)	6038.04	***					8444.78	***				

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table X.2: Mincer Wage Equation: Two-Step Heckman by Gender, 2014

	Female						Male					
	Probit			OLS			Probit			OLS		
2014: Gender	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>t	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>t
HH Size	-0.009	0.006					-0.025	0.006	***			
Employment	0.000	(omitted)					0.000	(omitted)				
Years of Schooling	0.000	0.008		0.103	0.007	***	0.012	0.008		0.091	0.005	***
Experience	0.021	0.008	***	0.024	0.007	***	0.026	0.008	***	0.019	0.005	***
Experience squared	-0.000	0.000	**	-0.000	0.000	**	-0.000	0.000	**	-0.000	0.000	**
Married and cohabitate	-0.406	0.079	***	-0.114	0.056	**	0.290	0.061	***	0.098	0.042	**
Other (Separated, divorced and widowed)	0.042	0.108		0.020	0.083		0.132	0.132		0.152	0.092	*
Religion1: Islam	-0.017	0.168		0.108	0.147		-0.135	0.184		0.028	0.114	
Religion2: Christian/Protestant	0.172	0.256		0.368	0.196	*	-0.296	0.259		0.198	0.156	
Religion3: Catholic	-0.003	0.208		0.393	0.179	**	-0.371	0.219	*	0.102	0.141	
Religion5: Buddhist	-0.019	0.804		0.119	0.542		0.433	0.628		0.037	0.358	
Ethnicity1: Jawa	-0.032	0.076		0.160	0.063	**	0.106	0.069		0.107	0.045	**
Ethnicity2: Sunda	-0.055	0.103		0.030	0.081		0.243	0.096	**	0.008	0.056	
Ethnicity3: Batak	-0.497	0.147	***	-0.044	0.128		-0.082	0.139		0.020	0.094	
Ethnicity4: Betawi	0.149	0.427		-0.243	0.264		0.015	0.306		-0.000	0.213	
Ethnicity5: Minang	0.025	0.134		0.158	0.121		-0.183	0.133		0.112	0.089	
Ethnicity6: Tiong Hoa	-0.591	0.338	*	0.281	0.298		-0.358	0.310		0.372	0.200	*
Status: fulltime (30 hours a week or more)	0.271	0.044	***	-0.320	0.045	***	0.401	0.049	***	-0.699	0.045	***
Tenure	-0.047	0.007	***	0.020	0.008	***	-0.006	0.007		0.015	0.005	***
Tenure squared	0.001	0.000	***	0.000	0.000		-0.000	0.000		0.000	0.000	
Sector: private	0.571	0.021	***	-0.274	0.059	***	1.101	0.022	***	-0.394	0.058	***
Industry2: mining and quarrying	0.196	0.487		0.599	0.369		0.586	0.157	***	0.310	0.095	***
Industry3: manufacturing	1.011	0.080	***	-0.020	0.110		0.843	0.076	***	0.007	0.068	
Industry4: electricity, gas and water	0.751	0.531		-0.137	0.488		0.872	0.222	***	0.101	0.121	
Industry5: construction	1.442	0.309	***	0.653	0.241	***	0.417	0.076	***	0.218	0.071	***
Industry6: wholesale, retail, restaurants and hotels	0.466	0.066	***	-0.139	0.096		0.348	0.064	***	-0.007	0.060	
Industry7: transportation, storage, and communications	0.831	0.319	***	0.328	0.240		0.650	0.105	***	0.023	0.084	
Industry8: Finance, insurance, real estate and business services	1.321	0.160	***	0.336	0.139	**	1.026	0.120	***	0.116	0.079	
Industry9: Social services	1.882	0.070	***	-0.020	0.138		1.216	0.064	***	-0.249	0.077	***
Firm size2: 5-19 people	1.406	0.051	***	0.226	0.096	**	0.956	0.048	***	-0.094	0.063	
Firm size3: 20-99 people	1.899	0.079	***	0.526	0.112	***	1.499	0.067	***	-0.025	0.077	
Firm size4: >= 100 people	2.349	0.116	***	0.902	0.125	***	1.443	0.082	***	0.224	0.077	***
Province 1: Aceh (11)	0.000	(omitted)		0.000	(omitted)		0.402	4.196		1.334	0.885	

Province 2: Sumatera Utara (12)	0.214	0.134		0.440	0.122	***	0.136	0.126		0.392	0.087	***
Province 3: Sumatera Barat (13)	0.346	0.169	**	0.624	0.152	***	0.459	0.165	***	0.221	0.113	**
Province 4: Riau (14)	-0.000	0.299		0.904	0.280	***	1.029	0.251	***	0.588	0.150	***
Province 5: Jambi (15)	-0.191	0.628		0.668	0.553		0.014	0.393		0.400	0.300	
Province 6: Sumatera Selatan (16)	0.170	0.124		0.491	0.123	***	0.379	0.112	***	0.248	0.081	***
Province 7: Lampung (18)	0.251	0.130	*	0.105	0.138		0.058	0.125		0.171	0.091	*
Province 8: Kepulauan Bangka Belitung (19)	0.573	0.281	**	0.746	0.186	***	-0.032	0.203		0.507	0.135	***
Province 9: Kepulauan Riau (21)	0.167	0.542		1.115	0.400	***	0.490	0.447		0.808	0.203	***
Province 10: DKI Jakarta (31)	0.493	0.126	***	0.829	0.103	***	0.022	0.113		0.615	0.070	***
Province 11: Jawa Barat (32)	0.246	0.123	**	0.531	0.105	***	-0.045	0.115		0.369	0.074	***
Province 12: Jawa Tengah (33)	0.351	0.118	***	0.124	0.103		-0.025	0.111		0.113	0.075	
Province 13: D I Yogyakarta (34)	0.132	0.138		0.158	0.115		0.011	0.136		-0.013	0.085	
Province 14: Jawa Timur (35)	0.107	0.110		0.166	0.099	*	-0.243	0.102	**	0.155	0.072	**
Province 15: Banten (36)	0.166	0.156		0.891	0.113	***	-0.045	0.140		0.454	0.084	***
Province 16: Bali (51)	0.339	0.179	*	0.872	0.155	***	0.142	0.193		0.433	0.118	***
Province 19: Kalimantan Tengah (62)	-0.223	0.510		0.865	0.435	**	0.103	0.497		1.419	0.286	***
Province 20: Kalimantan Selatan (63)	0.231	0.145		0.352	0.140	**	0.488	0.153	***	0.413	0.105	***
Province 21: Kalimantan Timur (64)	0.655	0.358	*	0.976	0.267	***	0.216	0.318		0.927	0.167	***
Province 22: Sulawesi Selatan (73)	0.050	0.120		0.192	0.107	*	-0.036	0.117		0.459	0.081	***
Province 23: Sulawesi Tenggara (74)	0.007	0.446		-0.265	0.548		0.363	0.444		0.256	0.298	
Urban	0.285	0.049	***	0.205	0.042	***	0.273	0.045	***	0.084	0.032	***
Mills				0.321	0.126	**				-0.465	0.134	***
Constants	-3.184	0.241	***	6.944	0.370	***	-4.501	0.245	***	8.925	0.344	***
Number of obs (Uncensored obs)	8170			3171			10623			4944		
The R-squared statistic				0.338						0.287		
LR chi2(50)	6396.32	***					9279.12	***				

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Appendix XI: Determinants of Education Mismatch in Selected Previous Empirical Studies

Table XI.1: Determinants of Education Mismatch

No	Author (s)	Countries	Study	Data	Education level	Method	Personal characteristics					Household Characteristics	
							Sex	Marital Status	Ethnicit y	Country of Birth	Year (cohort)	Young Children	Parent education
1	Battu and Sloane (2002)	UK	Determinants	the Fourth National Survey of Ethnic Minorities (FNSEM) conducted in 1993/94	Various education levels	Multinomia l logit	✓	✓	✓	✓			
2	Battu and Zakariya (2015)	Malaysia	Determinants and wage	the 2007 Productivity Investment Climate Survey	Various education levels	Multinomia l logit	✓	✓				✓	
3	Boll et al (2016)	EU	Determinants	the European Labour Force Survey (EU-LFS 2013)	High school and university	Probit	✓	✓	✓			✓	
4	Caroleo and Pastore (2013)	Italy	Determinants and wage	Alma Laurea university graduates, 2005	University	Logit	✓	✓	✓			✓	
5	Clark et al (2012)	US	Determinants and dynamic	the National Lon-gitudinal Survey of Youth 1979 (NLSY79) and CPS	Various education levels	Probit	✓		✓	✓			✓
6	Chevalier and Lindley (2007)	UK	Determinants and wage	UN Higher Education Institutions in 1995, primary survey	University	Multinomia l logit	✓		✓				
7	Flisi et al (2014)	EU	Determinants and wage	PIAAC data	Various education levels	Multinomia l logit	✓						
8	Green et all (1999)	UK	Determinants and wage	the National Child Development Study	Various education levels	Probit	✓				✓		
9	Linsley (2005)	Australia	Determinants	the Negotiating the Life Course survey, 1997	Various education levels	Probit		✓		✓		✓	

10	Linsley (2005)	Swiss	Determinants and wage	the Swiss Federal Statistical Office's (BFS) graduate survey, 2004-2011	University	Multinomial logit	✓						✓
11	McGuinness (2006)	Northern Ireland	Determinants and wage	Northern Ireland University graduates collected in 1999	University	Cohort		✓					
12	McGuinness et al (2017)	EU	Trends, convergence and drivers	the European Union Labour Force Study (EU-LFS) for the period Q1 1998 up to Q4 2012	Various education levels	Logit with Mundlak Correction	✓						
13	Meroni and Vera-Toscano (2017)	EU	Determinant and persistence	the 2005 REFLEX data (research into employment and professional flexibility)	Various education levels		✓						✓
14	Morano (2014)	Italy	Determinants	the National Labour Force Survey during the 2006-2011	Various education levels	Probit	✓	✓					
15	Kiker et al (1997)	Portugal	Determinants and wage		Years of education	Multinomial logit	✓						
16	Kiersztyn (2013)	Poland	Determinants and wage		High school and university	Logit	✓				✓		
17	Leuven and Oosterbeek (2011)		Literature review			Probit	✓		✓				
18	Raita (2005)	Dutch	Determinants and wage	NLSY 1979-2000	University	Probit	✓					✓	
19	Silles and Dolton (2002)	UK	Determinants	Newcastle alumni survey, 1998	University	Probit	✓	✓				✓	
20	Verhaest and Omey (2010)	Belgium	Determinants and wage	SONAR database, 1999	Various education levels	Probit	✓	✓	✓				
21	Yin (2016)	China	Determinants and wage	CHNS data from 1989 to 2009	Various education levels	Probit	✓				✓		

(cont'd)

No	Author (s)	A level score or other score	Education Level	Foreign Qualification	University Performance/degree	Education required	Subject	Study Grant	Skills/Ability	Age	Agesq
1	Battu and Sloane (2002)			✓					✓	✓	✓
2	Battu and Zakariya (2015)		✓								
3	Boll et al (2016)						✓			✓	
4	Caroleo and Pastore (2013)		✓		✓		✓				
5	Clark et al (2012)	✓								✓	
6	Chevalier and Lindley (2007)	✓	✓				✓		✓	✓	
7	Filsi et al (2014)		✓						✓	✓	
8	Green et all (1999)					✓				✓	
9	Linsley (2005)		✓							✓	
10	Diem (2015)			✓	✓					✓	
11	McGuinness (2006)	✓				✓	✓				
12	McGuinness et al (2017)										
13	Meroni and Vera-Toscano (2017)		✓				✓			✓	
14	Morano (2014)						✓			✓	
15	Kiker et al (1997)										
16	Kiersztyn (2013)		✓								
17	Leuven and Oosterbeek (2011)								✓	✓	
18	Raita (2005)		✓								
19	Silles and Dolton (2002)		✓		✓		✓			✓	
20	Verhaest and Omey (2010)				✓			✓	✓		
21	Yin (2016)		✓								

(cont’d)

No	Author (s)	Experience	Experiecesq	Tenure	Tenuresq	Unemployment Experience	Other experience	Occupation	Status	Sector	Industry
1	Battu and Sloane (2002)							✓		✓	
2	Battu and Zakariya (2015)	✓									
3	Boll et al (2016)			✓	✓				✓		
4	Caroleo and Pastore (2013)										
5	Clark et al (2012)	✓		✓		✓			✓		
6	Chevalier and Lindley (2007)										
7	Filsi et al (2014)										
8	Green et all (1999)										✓
9	Linsley (2005)			✓					✓		
10	Diem (2015)										
11	McGuinness (2006)	✓	✓							✓	
12	McGuinness et al (2017)								✓		
13	Meroni and Vera-Toscano (2017)										
14	Morano (2014)							✓	✓		✓
15	Kiker et al (1997)	✓	✓	✓	✓						
16	Kiersztyn (2013)			✓				✓			
17	Leuven and Oosterbeek (2011)										
18	Raita (2005)	✓						✓	✓		✓
19	Silles and Dolton (2002)	✓	✓			✓	✓	✓	✓	✓	
20	Verhaest and Omey (2010)						✓				
21	Yin (2016)	✓								✓	

(cont'd)

No	Author (s)	Firm Size	Immigration	Other	Residence	Regional Unemployment
1	Battu and Sloane (2002)	✓		Trade Union		
2	Battu and Zakariya (2015)			Training		
3	Boll et al (2016)	✓		marginal employment, second job, participation in LL		
4	Caroleo and Pastore (2013)		✓			
5	Clark et al (2012)			Number of job	✓	✓
6	Chevalier and Lindley (2007)					
7	Filsi et al (2014)					
8	Green et all (1999)					
9	Linsley (2005)					
10	Diem (2015)			Working before/while studying		
11	McGuinness (2006)			Work location		
12	McGuinness et al (2017)		✓			✓
13	Meroni and Vera-Toscano (2017)		✓		✓	
14	Morano (2014)					
15	Kiker et al (1997)					
16	Kiersztyn (2013)					✓
17	Leuven and Oosterbeek (2011)					
18	Raita (2005)	✓		Number of job		
19	Silles and Dolton (2002)	✓	✓			
20	Verhaest and Omey (2010)					✓
21	Yin (2016)	✓			✓	

Source: Author's compilation.

Table XI.2: Summary of Determinants of Education Mismatch

No	Variable	Number of studies
	Personal characteristic variables	
1	Sex	19 out of 21 studies in Table XI.1
2	Marital status	50 per cent of the studies in Table XI.1
3	Ethnicity	30 per cent of the studies in Table XI.1
4	Age and the square of age	More than 50 per cent of the studies in Table XI.1
	Household characteristics	
5	The presence of young children	Around 30 per cent of studies in Table XI.1
	Work related and firm size	
6	Tenure and tenure squared	5 out of 21 studies in Table XI.1
7	Working status	Around 30 per cent of the studies in Table XI.1
8	Sector (private/public)	4 studies in Table XI.1
9	Industries	3 studies in Table XI.1
10	Firm size	5 out of 21 studies in Table XI.1
	Region or area	
11	Capital province	Harris Todaro model
12	Urban	3 studies in Table XI.1

Source: Author's compilation, based on Table XI.1.

Appendix XII: Sample Distribution based on Occupation, Category and Sub-category, 2000 and 2014

Table XII.1: Sample Distribution Based on Occupation Category, 2000 and 2014

Occupation (1 digit)		2000						2014					
		N	mean	mode	SD	min	max	N	mean	mode	SD	min	max
1	Professional/Technical	151	12.29	12	2.65	6	18	328	13.85	12	2.46	6	22
2	The other professionals	568	13.86	12	2.42	6	18	1001	15.39	16	2	6	22
3	Administrative/managerial	25	14.08	12	2.16	12	18	57	13.61	16	3.1	6	18
4	Clerical and related workers	575	12.38	12	2.61	6	18	1072	12.91	12	2.78	6	18
5	Sales workers	575	10.38	12	3.09	6	18	922	11.72	12	3	6	22
6	Services workers	1040	9.31	6	3.05	6	16	1516	10.99	12	3.1	6	22
7	Agriculture, animal husbandry, forestry workers, fisherman and hunters	970	7.33	6	2.32	6	18	715	9.08	6	3.16	6	18
8	Craft and related trade workers	742	9.03	6	3	6	16	949	10.09	12	2.98	6	22
9	Plant and machine operators and assemblers	351	9.28	12	2.76	6	16	514	10.72	12	2.63	6	18
10	Elementary occupation	1331	8.56	6	2.84	6	16	1394	10.03	12	3.12	6	22
11	Armed forces occupation	57	11.18	12	2.21	6	18	21	13.29	12	1.87	12	16
	Number of observations	6385						8489					

Source: The author’s calculation.

Table XII.2: Sample Distribution Based on Occupation Sub-Category, 2000 and 2014

Code	Occupation	2000						2014					
		N	Mean	Mode	SD	Min	Max	N	Mean	Mode	SD	Min	Max
	Professional/Technical												
1	Physical scientists and related technicians	2	11	12	7.07	6	16	2	9	12	4.24	6	12
2	Architects, engineers, technologists	8	12.13	12	3.72	6	18	11	14.55	12	2.88	9	18
3	Surveyors, draftsmen, engineering assistants	40	11.5	12	2.48	6	16	59	12.49	12	2.6	6	18
4	Aircraft and ship's officer	3	15.67	12	0.58	15	16						
5	Life scientists and related technicians	2	12	12	0	12	12	3	14	12	1.73	12	15
6	Physicians, medical assistants, dentists, dental assistants, pharmacists, nutritionists	16	12.75	12	2.72	6	16	33	13.85	12	2.27	9	16
7	Nurses, midwives, x-ray technicians, traditional medicine practitioners	69	12.13	12	2.31	6	16	216	14.21	12	2.27	6	22
8	Statisticians, mathematicians, system analysts and related technicians	11	15	12	2.05	12	18	4	15.25	12	2.5	12	18
	The other professionals												
11	Accountants and auditors	10	14.8	12	1.93	12	16	8	15.38	16	2.33	12	18
12	Jurists	7	14.86	12	1.95	12	16	15	13.8	16	2.57	9	18
13	Teacher	451	14.25	12	2.1	6	18	878	15.68	16	1.54	6	22
14	Workers in religion	26	12.62	12	2.55	6	16	41	13.44	16	3.71	6	18
15	Authors, critics, journalists, editors and related writers	7	11.86	12	3.58	6	16	6	13.83	16	2.04	12	16
16	Sculptors, painters, photographers and related creative artists	12	11.42	12	3.15	6	16	19	11.05	16	3.84	6	16
17	Composers, performing artists	29	11.21	12	3	6	16	13	12	16	2.12	9	15
18	Athletes, sportsmen and related workers	5	13.6	12	3.05	9	16	6	14	16	2.19	12	16
19	Professional and technical workers not elsewhere classified assigned	21	12.19	12	2.6	6	16	15	14.87	16	2.23	12	18

	Administrative/managerial												
20	Legislative officials and government administrators	6	15.5	12	1.97	12	18	6	14.67	16	2.07	12	16
21	Managers	17	13.82	12	2.1	12	18	42	13.5	16	3.33	6	18
22	Managers							1	12	16		12	12
29	Manager not elsewhere classified (mostly school principals)	2	12	12	0	12	12	8	13.63	16	2.72	9	16
	Clerical and related workers												
30	Clerical supervisors	78	11.96	12	3.01	6	16	70	13.26	12	2.69	6	18
31	Government executive of officials	126	12.66	12	2.56	6	16	199	14.09	12	2.75	6	18
32	Stenographers, typists and card tape-punching machine operators	14	13.57	12	2.24	9	16	39	13.33	12	2.49	6	18
33	Bookkeepers, cashiers, and related workers	122	12.51	12	2.47	6	18	305	12.87	12	2.52	6	18
34	Computing machine operators	8	12.88	12	2.3	9	16	18	13.78	12	2.51	9	16
35	Transport and communications supervisors	1	12	12		12	12						
36	Transport conductors	11	9.55	12	2.62	6	12	21	10.14	12	2.92	6	15
37	Mail distributors and related workers	22	11.23	12	3.02	6	16	46	10.96	12	2.67	6	16
38	Telephone and telegraph operators	33	11.91	12	2.01	6	16	20	11.6	12	2.11	6	16
39	Clerical and related workers not elsewhere classified	160	12.59	12	2.49	6	16	354	12.62	12	2.77	6	18
	Sales workers												
40	Managers (wholesale and retail trades)	1	15	12		15	15						
41	Working proprietors (wholesale and retail trades)	2	12	12	0	12	12	32	13.59	12	2.88	6	16
42	Sales supervisors and buyers	6	11	12	2.45	6	12	39	12.44	12	3.43	6	22
43	Technical salesman, commercial travellers, manufacturer's agents	30	13.07	12	2.39	9	16	56	12.88	12	2.12	6	16

44	Insurance, real estate, securities and business services salesman and auctioneers	69	13.51	12	2.32	6	18	147	13.97	12	2.11	9	18
45	Salesmen, shop assistants and related workers	459	9.72	12	2.85	6	16	624	10.89	12	2.85	6	18
49	Sales workers not elsewhere classified	8	9.75	12	3.49	6	15	24	13.08	12	2.87	6	16
	Services workers												
50	Managers (catering and lodging services)							1	9	12		9	9
51	Working proprietors (catering and lodging services)	167	9.84	6	2.98	6	16	266	11.48	12	2.7	6	16
52	Housekeeping and related service supervisors							5	12	12	0	12	12
53	Cooks, waiters, bartenders and related workers	43	8.44	6	2.72	6	15	179	10.03	12	3	6	16
54	Maids and related housekeeping service workers NEC	451	8.58	6	3.03	6	16	424	10.47	12	3.72	6	22
55	Building caretakers, char workers, cleaners and related workers	75	9.08	6	2.86	6	16	158	10.54	12	2.77	6	18
56	Launderers, dry-cleaners and pressers	1	6	6		6	6	4	9	12	2.45	6	12
57	Hairdressers, barbers, beauticians and related workers	15	9.27	6	3.03	6	16	25	10.24	12	2.82	6	16
58	Protective service workers	121	10.23	6	2.56	6	16	216	11.46	12	2.3	6	16
59	Service workers not elsewhere classified	167	10.44	6	3.08	6	16	238	12.04	12	2.85	6	18
	Agriculture, animal husbandry, forestry workers, fisherman and hunters												
60	Plantation managers and supervisors	5	13.2	6	4.55	6	18	8	13.88	6	2.03	12	16
61	Planters and farmers	1	6	6		6	6	3	8	6	3.46	6	12
62	Agricultural and animal husbandry workers	654	6.98	6	2.06	6	16	462	8.8	6	3.05	6	18
63	Forestry workers	175	8.31	6	2.72	6	16	99	9.25	6	2.79	6	18
64	Fishermen, hunters, and related workers	135	7.53	6	2.26	6	15	142	9.66	6	3.53	6	18
69	Agricultural worker not elsewhere classified							1	6	6		6	6

	Craft and related trade workers												
70	Production supervisors and general foremen	77	11.6	6	2.52	6	16	100	11.9	12	2.78	6	18
71	Miners, quarrymen, well drillers and related workers	42	8.6	6	2.84	6	16	36	10.22	12	3.83	6	18
72	Metal processers	13	8.77	6	2.86	6	12	17	10.82	12	2.72	6	16
73	Wood preparation workers and paper makers	22	7.64	6	2.57	6	12	47	9.21	12	2.87	6	16
74	Chemical processers and related workers	43	8.95	6	3.13	6	16	65	9.94	12	2.88	6	18
75	Spinners, weavers, knitters, dyers, and related workers	115	8.57	6	2.67	6	16	80	9.31	12	2.88	6	16
76	Tanners, fellmongers and pelt dressers							6	10.67	12	2.88	9	16
77	Food and beverage processors	232	9.15	6	3.17	6	16	328	10.43	12	3.1	6	22
78	Tobacco preparers and tobacco product makers	9	7	6	2.12	6	12	13	7.62	12	2.33	6	12
79	Tailors, dressmakers, sewer, upholsterers and related workers	189	8.5	6	2.68	6	16	257	9.44	12	2.49	6	16
	Plant and machine operators and assemblers												
80	Shoemakers and leather good makers	34	8.82	12	2.66	6	15	59	10.34	12	2.54	6	16
81	Cabinet makers and related wood makers	99	8.46	12	2.65	6	16	92	10.53	12	2.84	6	16
82	Stone cutters and carvers	4	7.5	12	3	6	12	34	11.15	12	2.22	6	16
83	Blacksmith, tool makers and machine tool operators	34	9.44	12	2.57	6	12	28	11.04	12	2.01	6	15
84	Machinery fitters, assemblers, repairers and precision instrument makers (except electrical)	78	10.09	12	2.58	6	16	164	11.12	12	2.31	6	16
85	Electrical fitters and related electrical and electronics workers	42	11.05	12	2.54	6	16	60	10.78	12	2.73	6	18
86	Broadcasting station, sound equipment operators and cinema projectionists	1	12	12		12	12	3	11.33	12	5.03	6	16
87	Plumbers, welders, sheet-metal and structural metal preparers and erectors	32	8.53	12	2.65	6	12	42	9.55	12	2.91	6	16

88	Jewellery and precious metal workers	4	11.25	12	1.5	9	12	10	12.3	12	3.5	6	16
89	Glass formers, potters and related workers	23	8.22	12	2.75	6	12	22	9.73	12	3.01	6	16
	Elementary occupation												
90	Rubber and plastics product makers	27	8.67	6	2.09	6	12	37	9.46	12	2.87	6	16
91	Paper board products makers	9	9.67	6	2.5	6	12	23	10.39	12	3	6	16
92	Printers and related workers	73	9.71	6	3.03	6	16	101	10.41	12	2.76	6	16
93	Painters	24	8.13	6	2.25	6	12	35	9.43	12	2.32	6	12
94	Production and related workers not elsewhere classified	48	8.31	6	2.64	6	15	62	9.69	12	3.11	6	16
95	Bricklayers, carpenters and other construction workers	519	7.99	6	2.69	6	16	425	9.49	12	3.06	6	18
97	Material handling and related equipment, operators dockers and freight handlers	137	8.86	6	3.14	6	16	249	10.39	12	3.19	6	18
98	Transport equipment operators	339	8.81	6	2.66	6	16	304	10.19	12	2.77	6	18
99	Laborers not elsewhere classified	155	9.13	6	3.25	6	16	158	10.68	12	3.9	6	22
	Armed forces occupation												
M1	Military and Police	18	10.83	12	2.33	6	15	9	12.78	12	1.56	12	16
M2	Military and Police	22	11.77	12	2.33	6	18	12	13.67	12	2.06	12	16
MM	Military and Police	17	10.76	12	1.86	6	12						
		6385						8489					

Source: The author's calculation.
Note: N is number of observations, SD: standard deviations.

Table XII.3: The List of Occupations That Merge and The Same as The Original Sub-Category

The new occupation category, merging some occupation code/categories	**Occupation Code with remain same sub-category**
OX, comprises of occupation codes: 1, 2, 4, 5, 6, 8	3,7
1X, comprises of occupation codes: 11, 12, 14, 15, 16, 17, 18, 19	13
2X, comprises of occupation codes: 20, 21, 22, 23, 29	None
3X, comprises of occupation codes: 32, 34, 35, 36, 37, 38	30, 31, 33, 39
4X, comprises of occupation codes: 40, 41, 42, 43, 49	44, 45
5X, comprises of occupation codes: 50, 52, 53, 55, 56, 57	51, 54, 55, 58, 59
6X, comprises of occupation codes: 60, 61, 64, 69	62, 63
7X, comprises of occupation codes: 72, 73, 76, 78	70, 71, 74, 75, 77, 79
8X, comprises of occupation codes: 82, 83, 86, 88, 89	80, 81, 84, 85, 87
9X, comprises of occupation codes: 90, 91, 93, 96	92, 94, 95, 97, 98, 99
M, comprises of occupation codes: M1, M2, MM	None

Source: The author's compilation.
Note: the occupation name sub-category is indicated in previous table (Table XII.2)

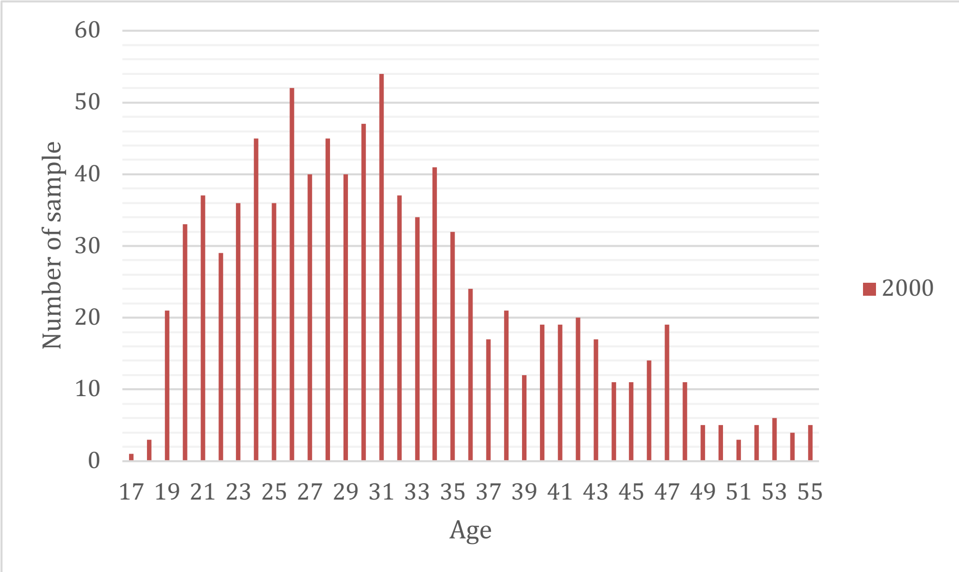
Table XII.4: Mismatch Distribution (1 Digit Occupation Category)

Year	Match (Mean)				Match (Mode)			
	UE	M	OE	Total	UE	M	OE	Total
All Sample								
2000								
Number of obs.	1,127	4,347	911	6,385	519	3,424	2,442	6,385
Proportion	17.65	68.08	14.27	100.00	8.13	53.63	38.25	100.00
2014								
Number of obs.	1,399	5,724	1,366	8,489	1,537	5,407	1,545	8,489
Proportion	16.48	67.43	16.09	100.00	18.11	63.69	18.20	100.00
Male								
2000								
Number of obs.	551	2,978	675	4,204	370	2,097	1,737	4,204
Proportion	13.11	70.84	16.06		8.8	49.88	41.32	
2014								
Number of obs.	795	3,494	866	5,155	899	3,315	941	5,155
Proportion	15.42	67.78	16.8		17.44	64.31	18.25	
Female								
2000								
Number of obs.	576	1,369	236	2,181	149	1,327	705	2,181
Proportion	26.41	62.77	10.82		6.83	60.84	32.32	
2014								
Number of obs.	604	2,230	500	3,334	638	2,092	604	3,334
Proportion	18.12	66.89	15		19.14	62.75	18.12	
Public								
2000								
Number of obs.	78	713	215	1,006	59	392	555	1,006
Proportion	7.75	70.87	21.37		5.86	38.97	55.17	
2014								
Number of obs.	67	882	456	1,405	73	829	503	1,405
Proportion	4.77	62.78	32.46		5.2	59	35.8	
Private								
2000								
Number of obs.	1,049	3,634	696	5,379	460	3,032	1,887	5,379
Proportion	19.5	67.56	12.94		8.55	56.37	35.08	
2014								
Number of obs.	1,332	4,842	910	7,084	1,464	4,578	1,042	7,084
Proportion	18.8	68.35	12.85		20.67	64.62	14.71	

Source: The author’s calculation.

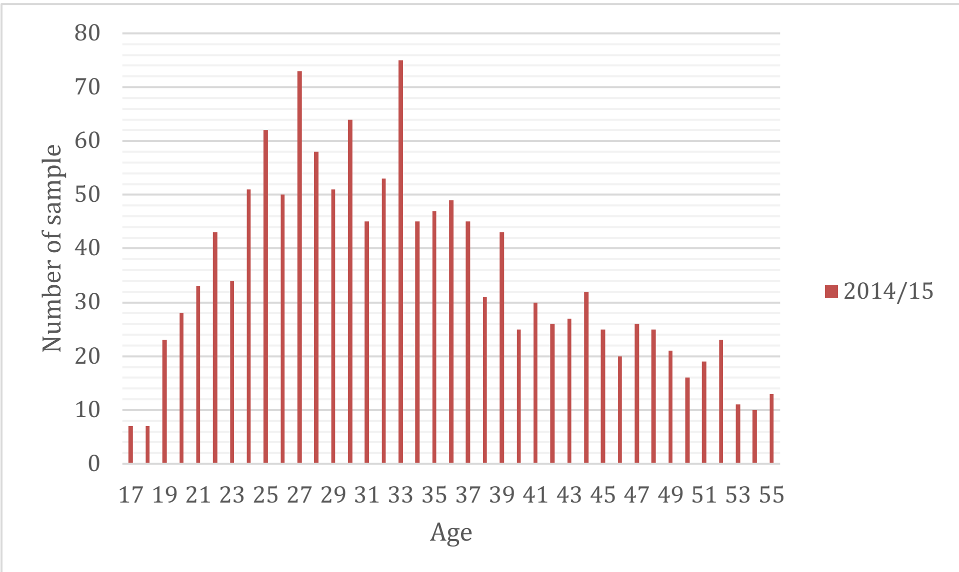
Appendix XIII: Sample Distributions: Age, Years of Schooling and Sex

Figure XIII.1: Age and Overeducation, 2000



Source: The author's calculation.

Figure XIII.2: Age and Overeducation, 2014



Source: The author's calculation.

Figure XIII.3: Years of Schooling and Sex

Years of schooling	Sex					
	2000			2014		
	Male	Female	Total	Male	Female	Total
0-6	1,485	871	2,356	878	579	1,457
	35.32	39.94	36.9	17.03	17.37	17.16
7-9	851	328	1,179	880	524	1,404
	20.24	15.04	18.47	17.07	15.72	16.54
10-12	1,394	618	2,012	2,269	1,079	3,348
	33.16	28.34	31.51	44.02	32.36	39.44
13-15	186	194	380	224	275	499
	4.42	8.9	5.95	4.35	8.25	5.88
13-16	275	169	444	820	807	1,627
	6.54	7.75	6.95	15.91	24.21	19.17
17-18	13	1	14	81	67	148
	0.31	0.05	0.22	1.57	2.01	1.74
19-22				3	3	6
				0.06	0.09	0.07
Total	4,204	2,181	6,385	5,155	3,334	8,489
	100	100	100	100	100	100

Source: The author’s calculation.

Appendix XIV: T-Test of Mismatch (Mode)

Figure XIV.1: T-Test of Mismatch

Mismatch (Mode)	Mean 2000	Mean 2014	Diff	T-Statistic	Degree of Freedom	P-value
Male						
UE	0.13	0.22	0.08	10.63	9357	***
M	0.55	0.53	-0.02	-2.05	9357	**
OE	0.32	0.26	-0.06	-6.76	9357	***
Female						
UE	0.14	0.24	0.10	9.25	5513	***
M	0.64	0.56	-0.08	-6.00	5513	***
OE	0.22	0.20	-0.02	-1.82	5513	*
Public						
UE	0.16	0.11	-0.05	-3.41	2409	***
M	0.55	0.57	0.02	1.07	2409	
OE	0.29	0.32	0.02	1.31	2409	
Private						
UE	0.13	0.25	0.12	16.57	12461	***
M	0.58	0.53	-0.05	-5.79	12461	***
OE	0.28	0.22	-0.07	-8.59	12461	***

Source: The author’s calculation.

Appendix XV: Summary Statistics with Casual Workers, 2014

Table XV.1: Summary Statistics of Sample with Casual Workers, 2014

Variable	Obs	Mean	SD	Min	Max
Match (Mode)	10594	2.03	0.68	1	3
Sex (1=Female)	10594	0.38	0.48	0	1
Marital Status	10594	1.83	0.48	1	3
Ethnicity (1=Javanese)	10594	0.46	0.50	0	1
Age	10594	33.71	9.78	16	55
Age squared	10594	1232.30	694.08	256	3025
Young children (0-5 years old)	10594	0.36	0.57	0	3
Tenure	10594	6.39	7.56	0	50
Tenure squared	10594	98.07	211.07	0	2500
Status: part time	10594	0.20	0.40	0	1
Sector: private	10594	0.67	0.47	0	1
Industry	10594	5.94	2.88	1	10
Firm Size	10594	1.57	0.77	1	3
Urban	10594	0.68	0.47	0	1
Capital	10594	0.08	0.27	0	1

Source: The author's calculation.

Appendix XVI: Sample Distribution and Summary Statistics of Education Mismatch by Gender and Sector

Table XVI.1: Sample Distribution Based on Years of Deficit Schooling and Gender (in per cent)

Years of Deficit Schooling	Male				Female			
	2000	2007	2014	Total	2000	2007	2014	Total
0	86.6	80.1	78.9	81.7	86.0	78.3	76.4	79.6
3	7.2	8.6	10.1	8.7	7.3	8.9	9.9	8.9
4	0.0	1.6	1.9	1.2	0.0	2.6	3.0	2.0
6	6.1	9.1	8.4	7.9	6.7	9.4	10.1	8.9
7	0.0	0.4	0.3	0.2	0.0	0.5	0.4	0.3
9	0.1	0.1	0.0	0.1	0.0	0.0	0.1	0.0
10	0.0	0.2	0.3	0.2	0.0	0.3	0.2	0.2
Total	4,204	3,953	5,155	13,312	2,181	2,347	3,334	7,862

Source: The author's calculation.

Table XVI.2: Sample Distribution Based on Years of Surplus Schooling and Gender

Years of Surplus Schooling	Male				Female			
	2000	2007	2014	Total	2000	2007	2014	Total
0	68.1	72.3	74.4	71.8	78.0	79.6	80.0	79.3
2	0.0	0.3	0.6	0.3	0.0	0.3	1.1	0.6
3	12.4	7.5	6.8	8.8	8.8	6.6	5.0	6.6
4	3.4	6.4	7.1	5.7	3.7	5.1	6.4	5.2
5	0.0	0.1	0.0	0.0	7.6	4.9	4.4	5.4
6	13.7	10.1	8.4	10.6	0.0	0.1	0.0	0.0
7	0.0	0.1	0.0	0.0	1.2	1.1	1.2	1.2
9	0.9	0.8	0.5	0.7	0.6	2.2	1.7	1.5
10	1.4	2.4	2.0	1.9	0.0	0.1	0.2	0.1
12	0.0	0.1	0.2	0.1	0.0	0.0	0.1	0.0
Total	4,204	3,953	5,155	13,312	2,181	2,347	3,334	7,862

Source: The author's calculation.

Table XVI.3: Summary Statistics of Main and Control Variables by Gender

Variable	Male					Female				
	Obs	Mean	SD	Min	Max	Obs	Mean	SD	Min	Max
Wage (log of hourly wages)	13,312	8.39	1.76	0	16.94	7,862	8.11	1.84	0	14.20
Years of required schooling	13,312	10.11	3.33	6	16.00	7,862	10.96	3.54	6	16.00
Years of surplus schooling (OE)	13,312	1.41	2.50	0	12.00	7,862	1.02	2.26	0	16.00
Years of deficit schooling (UE)	13,312	0.82	1.87	0	10.00	7,862	0.93	1.96	0	10.00
Potential Experience	13,312	16.36	10.07	0	43.00	7,862	14.37	10.36	0	43.00
Experience squared	13,312	368.99	395.5	0	1849	7,862	314	380.9	0	1849
Sex (1=female)	13,312	0.00	0.00	0	0	7,862	1.00	0.00	1	1
Ethnicity (1=Javanese)	13,312	0.15	0.36	0	1	7,862	0.16	0.37	0	1
Marital status: Single	13,312	0.25	0.43	0	1	7,862	0.28	0.45	0	1
Marital status: Married and cohabitate	13,312	0.74	0.44	0	1	7,862	0.65	0.48	0	1
Marital status: Other (Separated, divorced and widowed)	13,312	0.02	0.14	0	1	7,862	0.08	0.27	0	1
Employment Status: full-time (30 hours a week or more)	13,312	0.88	0.32	0	1	7,862	0.78	0.42	0	1
Tenure	13,312	6.72	7.42	0	51.83	7,862	6.10	7.19	0	52.00
Tenure squared	13,312	100.22	198.3	0	2686.7	7,862	88.84	192.8	0	2704.0
Sector: private	13,312	0.84	0.37	0	1	7,862	0.81	0.39	0	1
Industry1: agriculture	13,312	0.11	0.32	0	1	7,862	0.07	0.26	0	1
Industry2: mining and quarrying	13,312	0.02	0.13	0	1	7,862	0.00	0.04	0	1
Industry3: manufacturing	13,312	0.20	0.40	0	1	7,862	0.24	0.43	0	1
Industry4: electricity, gas and water	13,312	0.01	0.11	0	1	7,862	0.00	0.04	0	1
Industry5: construction	13,312	0.09	0.28	0	1	7,862	0.01	0.10	0	1
Industry6: wholesale, retail, restaurants and hotels	13,312	0.14	0.34	0	1	7,862	0.19	0.39	0	1
Industry7: transportation, storage, and communications	13,312	0.06	0.23	0	1	7,862	0.01	0.08	0	1
Industry8: finance, insurance, real estate and business services	13,312	0.04	0.20	0	1	7,862	0.03	0.17	0	1
Industry9: social services	13,312	0.33	0.47	0	1	7,862	0.44	0.50	0	1
Firm size1: 1-19 people	13,312	0.60	0.49	0	1	7,862	0.60	0.49	0	1
Firm size2: 20-99 people	13,312	0.24	0.43	0	1	7,862	0.24	0.43	0	1
Firm size3: >= 100 people	13,312	0.16	0.36	0	1	7,862	0.17	0.37	0	1
Urban	13,312	0.67	0.47	0	1	7,862	0.71	0.45	0	1
Capital region	13,312	0.10	0.31	0	1	7,862	0.11	0.31	0	1

Source: The author’s calculation.

Table XVI.4: RE Model with the Interaction Variables Between Sex and Education Mismatches

	Model 1			Model 2			Model 3			Model 4		
Wage	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>z
Years of required schooling	0.097	0.008	***	0.109	0.007	***	0.109	0.007	***	0.111	0.008	***
Years of surplus schooling (OE)	0.092	0.008	***	0.094	0.009	***	0.104	0.008	***	0.103	0.008	***
Years of deficit schooling (UE)	-0.076	0.009	***	-0.089	0.008	***	-0.072	0.010	***	-0.089	0.008	***
Experience	0.036	0.006	***	0.034	0.006	***	0.035	0.006	***	0.034	0.006	***
Experience squared	-0.001	0.000	***	-0.001	0.000	***	-0.001	0.000	***	-0.001	0.000	***
Sex (1=Female)	-0.634	0.108	***	-0.308	0.036	***	-0.238	0.036	***	-0.243	0.109	**
Married and cohabitate	0.177	0.043	***	0.177	0.043	***	0.176	0.043	***	0.176	0.043	***
Other (Separated, divorced and widowed)	0.219	0.084	***	0.207	0.084	**	0.210	0.084	**	0.204	0.084	**
Ethnicity (1=Javanese)	-0.023	0.031		-0.022	0.031		-0.023	0.031		-0.023	0.031	
Status: full-time (30 hours a week or more)	-0.390	0.040	***	-0.391	0.040	***	-0.392	0.040	***	-0.389	0.040	***
Tenure	0.073	0.006	***	0.074	0.006	***	0.074	0.006	***	0.075	0.006	***
Tenure squared	-0.002	0.000	***	-0.002	0.000	***	-0.002	0.000	***	-0.002	0.000	***
Sector: private	-0.275	0.048	***	-0.282	0.048	***	-0.279	0.048	***	-0.282	0.048	***
Industry2: mining and quarrying	0.359	0.133	***	0.361	0.133	***	0.359	0.133	***	0.360	0.133	***
Industry3: manufacturing	0.112	0.062	*	0.096	0.062		0.098	0.062		0.094	0.062	
Industry4: electricity, gas and water	-0.329	0.162	**	-0.346	0.162	**	-0.344	0.162	**	-0.344	0.162	**
Industry5: construction	0.095	0.083		0.116	0.083		0.115	0.083		0.113	0.083	
Industry6: wholesale, retail, restaurants and hotels	-0.026	0.065		-0.041	0.065		-0.041	0.065		-0.043	0.066	
Industry7: transportation, storage, and communications	-0.177	0.095	*	-0.182	0.095	*	-0.188	0.095	**	-0.178	0.096	*
Industry8: Finance, insurance, real estate and business services	0.380	0.087	***	0.365	0.087	***	0.368	0.087	***	0.366	0.087	***
Industry9: Social services	-0.034	0.061		-0.044	0.061		-0.046	0.061		-0.044	0.062	
Firm size2: 20-99 people	0.214	0.035	***	0.212	0.035	***	0.216	0.035	***	0.213	0.035	***
Firm size3: >= 100 people	0.446	0.041	***	0.438	0.041	***	0.443	0.041	***	0.438	0.041	***
Urban	0.204	0.034	***	0.203	0.034	***	0.203	0.034	***	0.203	0.034	***
Capital region	0.233	0.053	***	0.231	0.053	***	0.230	0.053	***	0.232	0.053	***
Sex*Years of Schooling	0.031	0.009	***									
Sex*Years of Surplus Schooling				0.026	0.012	**						
Sex*Years of Deficit Schooling							-0.041	0.015	***			

Sex*Years of Required Schooling										-0.003	0.009	
Constants	6.872	0.121	***	6.773	0.117	***	6.744	0.117	***	6.749	0.123	***
Sigma U	1.165			1.164			1.165			1.165		
Sigma E	1.308			1.309			1.308			1.309		
Rho	0.443			0.442			0.442			0.442		
Wald chi2(26)	2121.73			2113.86			2117.49			2108.84		
Prob > chi2	0			0			0			0		
Number of obs	14,789			14,789			14,789			14,789		
Number of groups	11,963			11,963			11,963			11,963		

Source: The author’s calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table XVI.5: Sample Distribution Based on Years of Deficit Schooling and Sector (in per cent)

Years of Deficit Schooling	Private				Public			
	2000	2007	2014	Total	2000	2007	2014	Total
0	86.7	76.8	75.7	79.4	84.5	90.2	89.3	88.3
3	6.3	9.8	11.2	9.3	12.0	4.5	4.2	6.4
4	0.0	1.7	1.9	1.3	0.0	3.0	4.4	2.7
6	6.9	11.0	10.6	9.6	3.2	2.0	1.3	2.0
7	0.0	0.5	0.3	0.3	0.0	0.2	0.5	0.2
9	0.1	0.0	0.0	0.0	0.3	0.2	0.0	0.2
10	0.0	0.3	0.3	0.2	0.0	0.0	0.3	0.1
Total	5,379	5,049	7,084	17,512	1,006	1,251	1,405	3,662

Source: The author's calculation.

Table XVI.6: Sample Distribution Based on Years of Surplus Schooling and Sector

Years of Surplus Schooling	Private				Public			
	2000	2007	2014	Total	2000	2007	2014	Total
0	71.7	77.2	78.3	75.9	70.7	66.3	68.0	68.2
2	0.0	0.2	0.3	0.2	0.0	0.8	3.5	1.6
3	12.1	8.1	6.7	8.8	6.3	3.6	3.2	4.2
4	2.4	4.9	6.1	4.6	9.3	10.2	10.3	10.0
6	12.0	7.7	6.9	8.7	0.0	0.2	0.0	0.1
7	0.0	0.1	0.0	0.0	9.6	9.8	6.1	8.3
9	0.9	0.7	0.5	0.7	1.9	2.0	2.2	2.0
10	1.0	1.2	1.1	1.1	2.1	7.0	5.8	5.2
12	0.0	0.1	0.1	0.1	0.1	0.2	0.8	0.4
16	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0
Total	5,379	5,049	7,084	17,512	1,006	1,251	1,405	3,662

Source: The author's calculation.

Table XVI.7: Summary Statistics of Main Variables by Sector

Variable	Private					Public				
	Obs	Mean	SD	Min	Max	Obs	Mean	SD	Min	Max
Wage (log of hourly wages)	17,512	8.12	1.80	0	16.94	3,662	9.08	1.55	0	13.88
Years of required schooling	17,512	9.99	3.24	6	16.00	3,662	12.48	3.58	6	16.00
Years of surplus schooling (OE)	17,512	1.15	2.25	0	16.00	3,662	1.81	3.04	0	16.00
Years of deficit schooling (UE)	17,512	0.94	1.99	0	10.00	3,662	0.47	1.38	0	10.00
Potential Experience	17,512	15.01	10.17	0	43.00	3,662	18.53	9.98	0	43.00
Experience squared	17,512	328.86	388.75	0	1849.00	3,662	442.79	388.41	0	1849.00
Sex (1=female)	17,512	0.36	0.48	0	1	3,662	0.40	0.49	0	1
Ethnicity (1=Javanese)	17,512	0.16	0.37	0	1	3,662	0.12	0.33	0	1
Marital status: Single	17,512	0.29	0.45	0	1	3,662	0.12	0.32	0	1
Marital status: Married and cohabitate	17,512	0.67	0.47	0	1	3,662	0.85	0.36	0	1
Marital status: Other (Separated, divorced and widowed)	17,512	0.04	0.20	0	1	3,662	0.03	0.18	0	1
Employment Status: full-time (30 hours a week or more)	17,512	0.85	0.35	0	1	3,662	0.79	0.41	0	1
Tenure	17,512	5.34	6.29	0	52.00	3,662	12.00	9.26	0	40.00
Tenure squared	17,512	68.05	158.18	0	2704.00	3,662	229.64	285.70	0	1600.00
Sector: private	17,512	1.00	0.00	1	1	3,662	0.00	0.00	0	0
Industry1: agriculture	17,512	0.11	0.31	0	1	3,662	0.04	0.19	0	1
Industry2: mining and quarrying	17,512	0.01	0.11	0	1	3,662	0.01	0.09	0	1
Industry3: manufacturing	17,512	0.26	0.44	0	1	3,662	0.03	0.16	0	1
Industry4: electricity, gas and water	17,512	0.01	0.08	0	1	3,662	0.01	0.10	0	1
Industry5: construction	17,512	0.07	0.25	0	1	3,662	0.01	0.11	0	1
Industry6: wholesale, retail, restaurants and hotels	17,512	0.19	0.39	0	1	3,662	0.01	0.10	0	1
Industry7: transportation, storage, and communications	17,512	0.04	0.20	0	1	3,662	0.02	0.13	0	1
Industry8: finance, insurance, real estate and business services	17,512	0.04	0.20	0	1	3,662	0.02	0.14	0	1
Industry9: social services	17,512	0.27	0.44	0	1	3,662	0.86	0.35	0	1
Firm size1: 1-19 people	17,512	0.62	0.49	0	1	3,662	0.50	0.50	0	1
Firm size2: 20-99 people	17,512	0.22	0.41	0	1	3,662	0.36	0.48	0	1
Firm size3: >= 100 people	17,512	0.16	0.37	0	1	3,662	0.14	0.35	0	1
Urban	17,512	0.69	0.46	0	1	3,662	0.65	0.48	0	1
Capital region	17,512	0.12	0.32	0	1	3,662	0.05	0.22	0	1

Source: The author’s calculation.

Appendix XVII: Estimation Results: Basic Model of Pooled OLS and Cluster Standard Error

Table XVII.1: Pooled OLS and Cluster Standard Error Results

Variable	Pool OLS			Cluster SE (Year)			Cluster SE (individual)		
	Coef	SD	P>t	Coef	SD	P>t	Coef	SD	P>t
Years of required schooling	0.120	0.005	***	0.106	0.003	***	0.106	0.005	***
Years of surplus schooling (OE)	0.108	0.006	***	0.100	0.004	**	0.100	0.006	***
Years of deficit schooling (UE)	-0.090	0.007	***	-0.087	0.004	**	-0.087	0.007	***
Experience	0.034	0.005	***	0.034	0.003	**	0.034	0.005	***
Experience squared	-0.001	0.000	***	-0.001	0.000	*	-0.001	0.000	***
Sex (1=female)	-0.305	0.025	***	-0.310	0.032	*	-0.310	0.025	***
Ethnicity (1=Javanese)	0.074	0.031	*	0.043	0.048		0.043	0.031	
Married and cohabitate	0.189	0.034	***	0.171	0.014	**	0.171	0.038	***
Other (Separated, divorced and widowed)	0.116	0.067		0.101	0.083		0.101	0.069	
Status: full-time (30 hours a week or more)	-0.367	0.033	***	-0.355	0.047	*	-0.355	0.040	***
Tenure	0.070	0.005	***	0.073	0.005	**	0.073	0.005	***
Tenure squared	-0.002	0.000	***	-0.002	0.000	**	-0.002	0.000	***
Sector: private	-0.296	0.038	***	-0.327	0.012	**	-0.327	0.038	***
Industry2: mining and quarrying	0.529	0.110	***	0.490	0.095	*	0.490	0.120	***
Industry3: manufacturing	0.140	0.047	**	0.153	0.061		0.153	0.048	**
Industry4: electricity, gas and water	-0.168	0.135		-0.171	0.289		-0.171	0.193	
Industry5: construction	0.161	0.060	**	0.166	0.029	*	0.166	0.068	*
Industry6: wholesale, retail, restaurants and hotels	0.009	0.051		0.008	0.060		0.008	0.054	
Industry7: transportation, storage, and communications	-0.020	0.070		0.008	0.166		0.008	0.078	
Industry8: finance, insurance, real estate and business services	0.440	0.074	***	0.391	0.089	*	0.391	0.074	***
Industry9: social services	0.040	0.046		0.057	0.089		0.057	0.048	
Firm size2: 20-99 people	0.229	0.028	***	0.219	0.026	*	0.219	0.027	***
Firm size3: >= 100 people	0.465	0.034	***	0.422	0.036	**	0.422	0.034	***
Urban	0.139	0.027	***	0.132	0.067		0.132	0.027	***
Capital region	0.227	0.038	***	0.260	0.022	**	0.260	0.042	***
2007				0.0848	0.003	***	0.085	0.029	**
2014				0.2923	0.005	***	0.292	0.031	***
Constants	6.570	0.087	***	6.619	0.050		6.619	0.097	***
Number of observations	21,174			21,174			21,174		
The R-squared statistic	0.15			0.15			0.16		
Coefficient test:	F-test (p-value)								
H0: years of required = years of OE	5.36 (0.02)								
H0: years of required = years of UE	461.76 (0.00)								

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Appendix XVIII: Coefficient Test by Gender and Sector

Table XVIII.1: Coefficient Test by Gender

	Male		
	F-test (p-value)		
	Pooled OLS + dummy year	RE	FE
H0: years of required = years of surplus schooling	2.23 (0.13)	0.47 (0.49)	1.79 (0.18)
H0: years of required = years of deficit schooling	209.80 (0.00)	171.41 (0.00)	3.17 (0.08)

	Female		
	F-test (p-value)		
	Pooled OLS + dummy year	RE	FE
H0: years of required = years of surplus schooling	0.25 (0.62)	0.82 (0.36)	0.25 (0.62)
H0: years of required = years of deficit schooling	142.58 (0.00)	121.62 (0.00)	2.17 (0.14)

Source: The author's calculation.

Table XVIII.1: Coefficient Test by Sector

	Private		
	F-test (p-value)		
	Pooled OLS + dummy year	RE	FE
H0: years of required = years of surplus schooling	0.00 (0.97)	0.02 (0.87)	0.15 (0.70)
H0: years of required = years of deficit schooling	219.82 (0.00)	193.10 (0.00)	3.61 (0.06)

	Public		
	F-test (p-value)		
	Pooled OLS + dummy year	RE	FE
H0: years of required = years of surplus schooling	0.23 (0.63)	0.30 (0.58)	3.15 (0.07)
H0: years of required = years of deficit schooling	129.29 (0.00)	102.27 (0.00)	0.01 (0.93)

Source: The author's calculation.

Appendix XIX: Robustness Test by Mean: Summary Statistics and Estimation Result

Table XIX.1: Summary Statistics of Selected Variables Based on Mean

Variable	Obs	Mean	SD	Min	Max
Years of required	21,174	10.72	2.20	6.98	15.64
Years of over-education	21,174	0.71	1.71	0.00	11.93
Years of under-education	21,174	0.62	1.50	0.00	9.64

Source: The author's calculation.

Table XIX.2: Coefficient Test

	Coefficient test		
	F-test (p-value)		
	Pooled OLS	**RE**	**FE**
H0: years of required = years of OE	17.03 (0.00)	12.67 (0.00)	0.92 (0.34)
H0: years of required = years of UE	322.09 (0.00)	252.25 (0.00)	0.93 (0.34)

Source: The author's calculation.
Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Table XIX.3: Estimation Results Based on Mean

Variable	Pooled OLS with dummy year			RE			FE		
	Coef	SD	P>t	Coef	SD	P>t	Coef	SD	P>t
Years of required schooling	0.133	0.007	***	0.121	0.008	***	0.012	0.019	
Years of surplus schooling (OE)	0.096	0.007	***	0.090	0.007	***	-0.002	0.018	
Years of deficit schooling (UE)	-0.076	0.008	***	-0.073	0.009	***	-0.019	0.019	
Experience	0.029	0.005	***	0.026	0.005	***	-0.018	0.020	
Experience squared	-0.001	0.000	***	-0.001	0.000	***	0.000	0.000	
Sex (1=female)	-0.340	0.025	***	-0.355	0.028	***	(omitted)		
Ethnicity (1=Javanese)	0.038	0.032		0.053	0.036		(omitted)		
Married and cohabitate	0.190	0.034	***	0.194	0.035	***	0.238	0.070	***
Other (Separated, divorced and widowed)	0.117	0.066		0.135	0.068	*	0.258	0.130	*
Status: full-time (30 hours a week or more)	-0.344	0.033	***	-0.359	0.033	***	-0.440	0.058	***
Tenure	0.074	0.005	***	0.070	0.005	***	0.035	0.008	***
Tenure squared	-0.002	0.000	***	-0.002	0.000	***	-0.001	0.000	***
Sector: private	-0.340	0.038	***	-0.331	0.040	***	-0.217	0.087	*
Industry2: mining and quarrying	0.465	0.110	***	0.372	0.113	***	0.088	0.206	
Industry3: manufacturing	0.155	0.047	***	0.181	0.047	***	0.137	0.092	
Industry4: electricity, gas and water	-0.187	0.135		-0.179	0.137		-0.303	0.241	
Industry5: construction	0.142	0.060	*	0.144	0.061	*	0.079	0.119	
Industry6: wholesale, retail, restaurants and hotels	0.030	0.050		0.051	0.051		0.015	0.101	
Industry7: transportation, storage, and communications	0.030	0.070		0.054	0.071		0.015	0.130	
Industry8: Finance, insurance, real estate and business services	0.361	0.075	***	0.361	0.075	***	0.161	0.143	
Industry9: Social services	0.012	0.048		0.050	0.048		0.009	0.093	
Firm size2: 20-99 people	0.229	0.028	***	0.215	0.029	***	0.069	0.050	
Firm size3: >= 100 people	0.454	0.035	***	0.416	0.035	***	0.133	0.063	*
Urban	0.153	0.027	***	0.168	0.028	***	0.091	0.067	
Capital region	0.259	0.039	***	0.256	0.043	***	0.288	0.197	
2007	0.104	0.030	***	0.115	0.028	***	0.519	0.133	***
2014	0.299	0.030	***	0.335	0.029	***	1.138	0.251	***
Constants	6.372	0.100	***	6.484	0.104	***	8.035	0.309	***
Rho				0.438			0.652		
Number of observations	21,174			21,174			21,174		
Number of individuals				15,440			15,440		
The R-squared statistic	0.1557								
Wald chi2(3)				3298.850		***			
F Test							29.49		***
Post Estimation:									
Breusch and Pagan Lagrangian multiplier test for RE									
Var(u) = 0				29.840		***			
Hausman Test									
H0: difference in coefficients not systematic							140.9		***

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Appendix XX: Robustness test of the Balanced Panel Data: Sample Distribution, Summary Statistics and Estimation Result

Table XX.1: Distribution of Years of Required Schooling, Years of Surplus Schooling and Years of Deficit Schooling for Balanced Panel Data

Years of Required Schooling	**2000**	**2007**	**2014**	**Total**
6	468	275	205	948
9	13	29	0	42
12	464	567	649	1,680
15	183	42	26	251
16	0	215	248	463
Total	1,128	1,128	1,128	3,384

Years of Surplus Schooling	**2000**	**2007**	**2014**	**Total**
0	796	892	899	2,587
2	0	6	17	23
3	102	55	47	204
4	46	48	63	157
6	160	88	65	313
9	11	10	6	27
10	13	27	27	67
12	0	2	4	6
Total	1,128	1,128	1,128	3,384

Years of Deficit Schooling	**2000**	**2007**	**2014**	**Total**
0	967	842	812	2,621
3	112	118	119	349
4	0	36	38	74
6	49	122	153	324
7	0	6	5	11
9	0	2	0	2
10	0	2	1	3
Total	1,128	1,128	1,128	3,384

Source: The author’s calculation.

Table XX.2: Summary Statistics of the Balanced Panel

Variable	Obs	Mean	SD	Min	Max
Wage (log of hourly wages)	3,384	8.69	1.41	0	13.42
Years of required schooling	3,384	11.05	3.49	6	16
Years of surplus schooling (OE)	3,384	1.23	2.45	0	12
Years of deficit schooling (UE)	3,384	1.01	2.00	0	10
Experience	3,384	19.46	8.68	0	43
Experience squared	3,384	453.98	357.14	0	1849
Sex (1=female)	3,384	0.33	0.47	0	1
Marital status: Single	3,384	0.12	0.33	0	1
Marital status: Married and cohabitate	3,384	0.84	0.37	0	1
Marital status: Other (Separated, divorced and widowed)	3,384	0.04	0.20	0	1
Status: full-time (30 hours a week or more)	3,384	0.33	0.47	0	1
Tenure	3,384	0.86	0.34	0	1
Tenure squared	3,384	10.21	8.41	0	51.83333
Sector: private	3,384	174.85	240.48	0	2686.694
Industry1: agriculture	3,384	0.06	0.24	0	1
Industry2: mining and quarrying	3,384	0.01	0.10	0	1
Industry3: manufacturing	3,384	0.23	0.42	0	1
Industry4: electricity, gas and water	3,384	0.01	0.09	0	1
Industry5: construction	3,384	0.04	0.21	0	1
Industry6: wholesale, retail, restaurants and hotels	3,384	0.12	0.32	0	1
Industry7: transportation, storage, and communications	3,384	0.04	0.19	0	1
Industry8: finance, insurance, real estate and business services	3,384	0.03	0.16	0	1
Industry9: Social services	3,384	0.46	0.50	0	1
Firm size1: 1-19 people	3,384	0.53	0.50	0	1
Firm size2: 20-99 people	3,384	0.29	0.45	0	1
Firm size3: >= 100 people	3,384	0.18	0.38	0	1
Urban	3,384	0.71	0.45	0	1
Capital region	3,384	0.09	0.29	0	1

Source: The author’s calculation.

Table XX.3: Estimation Results of the Balanced Panel Data

Variable	Pooled OLS with dummy year			RE			FE		
	Coef	SD	P>t	Coef	SD	P>t	Coef	SD	P>t
Years of required schooling	0.119	0.010	***	0.119	0.010	***	0.080	0.032	*
Years of surplus schooling (OE)	0.105	0.011	***	0.107	0.012	***	0.084	0.032	**
Years of deficit schooling (UE)	-0.113	0.012	***	-0.116	0.013	***	-0.092	0.034	**
Experience	0.013	0.012		0.013	0.012		-0.012	0.034	
Experience squared	0.000	0.000		0.000	0.000		0.000	0.000	
Sex (1=female)	-0.194	0.047	***	-0.194	0.050	***		(omitted)	
Ethnicity (1=Javanese)	-0.071	0.052		-0.070	0.056			(omitted)	
Married and cohabitate	0.265	0.081	**	0.260	0.083	**	0.198	0.110	
Other (Separated, divorced and widowed)	0.173	0.132		0.197	0.134		0.375	0.190	*
Status: full-time (30 hours a week or more)	-0.333	0.065	***	-0.341	0.065	***	-0.399	0.086	***
Tenure	0.029	0.008	***	0.030	0.008	***	0.030	0.010	**
Tenure squared	0.000	0.000		0.000	0.000		-0.001	0.000	*
Sector: private	-0.512	0.063	***	-0.491	0.066	***	-0.181	0.119	
Industry2: mining and quarrying	0.753	0.218	***	0.695	0.223	**	0.162	0.314	
Industry3: manufacturing	0.146	0.096		0.156	0.098		0.189	0.134	
Industry4: electricity, gas and water	-0.572	0.239	*	-0.573	0.244	*	-0.667	0.341	
Industry5: construction	0.086	0.130		0.104	0.132		0.287	0.187	
Industry6: wholesale, retail, restaurants and hotels	-0.102	0.109		-0.084	0.110		0.041	0.153	
Industry7: transportation, storage, and communications	-0.009	0.136		0.035	0.139		0.404	0.192	*
Industry8: finance, insurance, real estate and business services	0.175	0.156		0.158	0.159		-0.002	0.220	
Industry9: Social services	-0.117	0.098		-0.087	0.099		0.075	0.134	
Firm size2: 20-99 people	0.110	0.049	*	0.103	0.051	*	0.001	0.074	
Firm size3: >= 100 people	0.146	0.062	*	0.136	0.064	*	-0.004	0.091	
Urban	0.010	0.049		0.009	0.052		-0.021	0.098	
Capital region	0.325	0.075	***	0.318	0.080	***	-0.105	0.344	
2007	0.106	0.059		0.109	0.058		0.401	0.233	
2014	0.321	0.072	***	0.326	0.073	***	0.856	0.439	
Constants	7.172	0.187	***	7.148	0.194	***	7.673	0.705	***
Rho				0.084			0.352		
Number of observations	3,384			3,384			3,384		
Number of individuals				1,128			1,128		
The R-squared statistic	0.2681								
Wald chi2(3)				1122.9		***			
F Test							15.01		***
Post Estimation:									
Breusch and Pagan Lagrangian multiplier test for RE									
Var(u) = 0				21.690		***			
Hausman Test									
Ho: difference in coefficients not systematic							71.3		***

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.

Appendix XXI: Estimation Results Based on Verdugo and Verdugo's Model

Table XXI.1: Verdugo and Verdugo's Estimation Results

Variable	Pooled OLS with dummy year			RE			RE		
	Coef	**SD**	**P>t**	**Coef**	**SD**	**P>t**	**Coef**	**SD**	**P>t**
Years of actual schooling	0.112	0.005	***	0.109	0.006	***	0.059	0.033	
Dummy of Overeducation	0.118	0.034	***	0.094	0.033	**	-0.041	0.055	
Dummy of Undereducation	-0.034	0.030		-0.016	0.030		0.026	0.051	
Experience	0.035	0.005	***	0.033	0.005	***	0.025	0.029	
Experience squared	-0.001	0.000	***	-0.001	0.000	***	0.000	0.000	*
Sex (1=female)	-0.313	0.025	***	-0.329	0.028	***		(omitted)	
Ethnicity (1=Javanese)	0.042	0.031		0.056	0.036			(omitted)	
Married and cohabitate	0.169	0.034	***	0.170	0.035	***	0.236	0.070	***
Other (Separated, divorced and widowed)	0.100	0.066		0.116	0.068		0.253	0.130	
Status: full-time (30 hours a week or more)	-0.356	0.033	***	-0.367	0.033	***	-0.441	0.058	***
Tenure	0.073	0.005	***	0.068	0.005	***	0.035	0.008	***
Tenure squared	-0.002	0.000	***	-0.002	0.000	***	-0.001	0.000	***
Sector: private	-0.318	0.038	***	-0.300	0.040	***	-0.216	0.087	*
Industry2: mining and quarrying	0.477	0.110	***	0.379	0.112	***	0.086	0.206	
Industry3: manufacturing	0.145	0.047	**	0.172	0.047	***	0.144	0.091	
Industry4: electricity, gas and water	-0.182	0.135		-0.177	0.137		-0.286	0.241	
Industry5: construction	0.168	0.060	**	0.162	0.061	**	0.075	0.119	
Industry6: wholesale, retail, restaurants and hotels	-0.003	0.051		0.022	0.052		0.027	0.101	
Industry7: transportation, storage, and communications	0.001	0.070		0.028	0.070		0.019	0.129	
Industry8: finance, insurance, real estate and business services	0.373	0.074	***	0.363	0.074	***	0.176	0.142	
Industry9: Social services	0.027	0.046		0.056	0.047		0.022	0.091	
Firm size2: 20-99 people	0.217	0.028	***	0.202	0.029	***	0.071	0.050	
Firm size3: >= 100 people	0.421	0.035	***	0.385	0.035	***	0.135	0.063	*
Urban	0.123	0.027	***	0.138	0.028	***	0.089	0.067	
Capital region	0.258	0.038	***	0.254	0.043	***	0.291	0.197	
2007	0.085	0.030	**	0.093	0.028	***	0.222	0.204	
2014	0.289	0.030	***	0.315	0.029	***	0.563	0.388	
Constants	6.548	0.088	***	6.564	0.092	***	7.177	0.556	***
Rho				0.437			0.628		
Number of observations	21,174			21,174			21,174		
Number of individuals				15,440			15,440		
The R-squared statistic	0.159								
Wald chi2(3)				3406.81		***			
F Test							29.64		***
Post Estimation:									
Breusch and Pagan Lagrangian multiplier test RE									
Var(u) = 0				30.860		***			
Hausman Test									
Ho: difference in coefficients not systematic							98.73		***

Source: The author's calculation.

Notes: * significance level at 10 per cent; ** significance level at 5 per cent; *** significance level at 1 per cent.